OPTIMAL TRANSPORT, CHEEGER ENERGIES AND CONTRACTIVITY OF DYNAMIC TRANSPORT DISTANCES IN EXTENDED SPACES

LUIGI AMBROSIO, MATTHIAS ERBAR, AND GIUSEPPE SAVARÉ

Dedicated to J.L. Vazquez in occasion of his 70th birthday

ABSTRACT. We introduce the setting of *extended metric-topological measure spaces* as a general "Wiener like" framework for optimal transport problems and nonsmooth metric analysis in infinite dimension.

After a brief review of optimal transport tools for general Radon measures, we discuss the notions of the Cheeger energy, of the Radon measures concentrated on absolutely continuous curves, and of the induced "dynamic transport distances". We study their main properties and their links with the theory of Dirichlet forms and the Bakry-Émery curvature condition, in particular concerning the contractivity properties and the EVI formulation of the induced Heat semigroup.

CONTENTS

1. Introduction

In the last years many papers have been devoted to the investigation of the connection between gradient contractivity, contractivity of transport distances and lower bounds on Ricci curvature and to the connection between metric and differentiable structures. In these investigations one can take as starting point either a metric measure space $(X, \mathsf{d}, \mathfrak{m})$ or a Dirichlet form $\mathcal{E}$ in $L^2(X, \mathfrak{m})$. In particular [5] provided key connections between the two viewpoints, proving that under mild regularity assumptions the distance $\mathsf{d}_\mathcal{E}$ generated out of the Dirichlet form as in [13] induces a metric energy (called Cheeger energy in [3], [4]) equal to $\mathcal{E}$, and that Ricci lower bounds can be equivalently stated either in terms of the Bakry-Émery gradient K-contractivity condition $\mathsf{BE}(K, \infty)$, $K \in \mathbb{R}$,

$$\Gamma(\mathsf{P}_t f) \leq \mathrm{e}^{-2Kt} \, \mathsf{P}_t \Gamma(f) \tag{1.1}$$

(here P is the semigroup induced by $\mathcal{E}$), or in terms of K-convexity of the entropy along Wasserstein geodesics (see also [30, 31] and also [22] for extensions to the case when upper bounds on the dimension are considered). The crucial link between the two formulations is provided by the characterization of the semigroup P as the EVI_K-gradient flow (see (1.2) below) of the entropy in the Wasserstein space.

A typical assumption made in the above-mentioned papers is that the topological/measure structure is induced by the distance, and that the distance is finite: for instance, when one takes $\mathcal{E}$ as starting point, one assumes that the topology induced by $\mathsf{d}_\mathcal{E}$ coincides with the initial topology of the space. However, there exist examples where the topology induced by the natural distance is too fine and the distance can be even infinite: the simplest and probably most studied and natural example is the so-called Wiener space, i.e. a Gaussian measure space endowed with the Cameron-Martin distance.

The main goal of this and of the forthcoming paper [8] is a deeper investigation of the above-mentioned problems in *extended* metric structures, where extended metric spaces are sets X endowed with a symmetric and triangular $\mathsf{d} : X \times X \to [0, \infty]$, with $\mathsf{d}(x, y) = 0$ iff $x = y$. Extended distances arise in a natural way either by taking the supremum $\sup |f(x) - f(y)|$ along a set $\mathcal{F}$ of functions which separate the points of X (this is precisely what happens with $\mathsf{d}_\mathcal{E}$), by construction of length distances and more generally by action minimization. At this level many extension of the classical metric theory, for instance the existence of metric derivatives $|\dot{x}|(t)$ for absolutely continuous curves $x(t)$ are fairly trivial, since d induces equivalence classes in X which are classical metric spaces; on the other hand, already the example of the Wiener space shows that when when we are given a reference measure $\mathfrak{m}$ on X it is very hard to work with the quotient structure, and it is much better to consider the space as a whole; also in many cases it happens that we are given a topology τ in X, coarser than the topology induced by the extended distance. We axiomatize this richer structure with the concept of *extended metric-topological space* (X, τ, d), characterized by the existence of a family $\mathcal{A}$ of bounded functions which separate the points of X and generate both the Hausdorff topology τ and the distance $\mathsf{d}(x, y)$, the latter with the formula $\sup_{f \in \mathcal{A}} |f(x) - f(y)|$. We denote in the sequel the algebra $\mathrm{Lip}_b(X, \tau, \mathsf{d})$ of bounded, d-Lipschitz and τ-continuous functions, which includes $\mathcal{A}$ and generates τ as well. In view of the applications we have in mind in [8], we are not assuming that the family $\mathcal{A}$ is countable and, correspondingly, we do not need extra assumptions on τ; the complete regularity of τ is implied by the $\mathcal{A}$-generating property and it will be sufficient for our purpose, provided the reference measure $\mathfrak{m}$ is Radon.

Now we pass to a more detailed description of the content of the paper. Section 2 is devoted to some measure-theoretic preliminaries, mostly borrowed from the very comprehensive monographs [14, 39]. In particular we introduce the class of Radon measures, denoted by $\mathscr{P}(X)$, and recall the basic compactness theorem for families of probability measures (see Theorem 2.1). Then we recall the dual formulation of the optimal transport problem when the marginals are Radon measures, following [29] (see also [50] for the analysis of optimal transport problem in a very general setup).

Section 3 contains basic and already well-established results of the metric theory, with their easy adaptation to the extended setting. The only, but essential, new ingredient is a self-improvement principle for solutions to the so-called EVI_K (evolution variational inequality) gradient flows along a semigroup S

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\frac{1}{2}\mathsf{d}^2(\mathsf{S}_t x, y) + \frac{K}{2}\mathsf{d}^2(\mathsf{S}_t x, y) \leq F(y) - F(\mathsf{S}_t x) \qquad \forall t > 0 \ , \tag{1.2}$$

(for all y at a finite d-distance from some $\mathsf{S}_t x$) which allows to some extent to pass from (1.2) to the same inequality for the (extended) length distances $\mathsf{d}_\ell, \bar{\mathsf{d}}_\ell$ induced by d (the former defined in (3.8) by ε-chains, the latter defined in (3.9) by the minimization of the length of curves), see Theorem 3.5 for a precise statement which involves a powerful integral formulation of (1.2). Moreover in Corollary 3.6 we derive from EVI_K a discrete convexity property relative to d_ℓ that, under suitable compactness assumptions, can be improved to convexity along all geodesics.

In Section 4 we introduce metric-topological spaces (X, τ, d) and we discuss a few preliminary properties of them, in particular the density of $\mathrm{Lip}_b(X, \tau, \mathsf{d})$ in $L^2(X, \mathfrak{m})$ for $\mathfrak{m}$ Radon, compactness of measures in the space X^D of paths and lower semicontinuity of the p-action, $p \in (1, \infty)$, in this context. In Section 5 we extend to this setting (a priori neither separable nor metrizable) basic results relative to the extended distance in $\mathscr{P}(X)$ induced by the quadratic cost d^2. Denoting this distance by W_d, we prove compactness and lower semicontinuity theorems, the implication from W_d-convergence to weak convergence and the basic superposition theorem which extends to a non-Polish setup the recent paper [34]. Thanks to this result, 2-absolutely continuous curves μ_t in $(\mathscr{P}(X), W_2)$ have a lifting to the space $X^{[0,1]}$, i.e. there exists $\boldsymbol{\eta} \in \mathscr{P}(X^{[0,1]})$ concentrated on 2-absolutely continuous paths $\eta : [0,1] \to (X, \mathsf{d})$ whose marginals are μ_t and satisfying

$$\int |\dot{\eta}(t)|^2 \,\mathrm{d}\boldsymbol{\eta}(\eta) = |\dot{\mu}_t|^2 \qquad \text{for a.e. } t \in (0,1) \ .$$

In Section 6 we recall the basic construction of the so-called Cheeger energy [17], adapted to the extended setting (X, τ, d) with a reference measure $\mathfrak{m} \in \mathscr{P}(X)$. Following with minor variants [3] (these variants allow to bypass some measurability issues relative to the asymptotic Lipschitz constant) we set

$$\mathrm{Ch}(f) := \inf \liminf_{n \to \infty} \int g_n^2 \,\mathrm{d}\mathfrak{m} \ ,$$

where the infimum runs among all sequences $(f_n) \subset \mathrm{Lip}_b(X, \tau, \mathsf{d})$ with $\lim_n \int |f_n - f|^2 \,\mathrm{d}\mathfrak{m} = 0$ and all $\mathfrak{m}$-measurable functions $g_n \geq \mathrm{Lip}_a(f_n, \cdot)$ $\mathfrak{m}$-a.e. in X (where Lip_a is the so-called asymptotic Lipschitz constant, see (6.1)). Together with the construction of Ch there is the construction of a local object, called minimal relaxed slope and denoted with $|\mathrm{D}f|_w$, which provides integral representation to Ch via $\mathrm{Ch}(f) = \int |\mathrm{D}f|_w^2 \,\mathrm{d}\mathfrak{m}$ when $\mathrm{Ch}(f)$ is finite. As shown in [17] and [3], many classical properties of Sobolev functions extend to this setting; in addition, defining Δf as the element with minimal norm in the subdifferential $\partial \mathrm{Ch}(f)$, it is well-defined a Heat flow in $L^2(X, \mathfrak{m})$ (linear iff Ch is quadratic) P, given by $\frac{d}{dt}\mathsf{P}_t f = \Delta \mathsf{P}_t f$, according to the evolution theory for maximal monotone operators in Hilbert spaces.

The aim of Section 7 is to introduce two more extended distances induced by Ch, in the subset $\mathscr{P}^a(X) \subset \mathscr{P}(X)$ of measure absolutely continuous w.r.t. $\mathfrak{m}$. The first one, denoted by W_{Ch}, is a length distance in $\mathscr{P}^a(X)$ whose definition is inspired by the Benamou-Brenier formula

$$W_{\mathrm{Ch}}^2(\rho_0 \mathfrak{m}, \rho_1 \mathfrak{m}) := \inf \left\{ \int_0^1 \|\rho_t'\|^2 \,\mathrm{d}t : \ \rho_t \in \mathsf{CE}^2(X, \mathrm{Ch}, \mathfrak{m}) \right\} \ .$$

Here $\|\rho_t'\|$ is the least function $c(t)$ in $L^2(0,1)$ satisfying

$$\left| \int f \rho_s \,\mathrm{d}\mathfrak{m} - \int f \rho_t \,\mathrm{d}\mathfrak{m} \right| \leq \int_s^t c(r) \left(\int |\mathrm{D}f|_w^2 \rho_r \,\mathrm{d}\mathfrak{m} \right)^{1/2} \,\mathrm{d}r \qquad \forall f \in \mathrm{Lip}_b(X, \tau, \mathsf{d})$$

and the property above defines the class of curves in $\mathsf{CE}^2(X, \mathrm{Ch}, \mathfrak{m})$. The second one, denoted $W_{\mathrm{Ch},*}$, has a dual character and it is defined by

$$W_{\mathrm{Ch},*}^2(\rho_0, \rho_1) := 2 \sup_{\phi} \int (\phi_1 \rho_1 - \phi_0 \rho_0) \, \mathrm{d}\mathfrak{m} \,,$$

where the supremum runs along all the "formal" subsolutions to the Hamilton-Jacobi equation $\frac{\mathrm{d}}{\mathrm{d}t}\phi_t + |\mathrm{D}\phi_t|_w^2/2 = 0$. At this high level of generality, we are only able to prove that $W_{\mathsf{d}} \leq W_{\mathrm{Ch},*} \leq W_{\mathrm{Ch}}$; however, when we pass from the global to the infinitesimal behaviour, these distances reveal much closer connections. Indeed, under the Bakry-Émery gradient contractivity condition $|\mathrm{D}\mathsf{P}_t f|_w^2 \leq \mathrm{e}^{-2Kt} \mathsf{P}_t |\mathrm{D}f|_w^2$ for some $K \in \mathbb{R}$ one can prove that the length distance associated to $W_{\mathrm{Ch},*}$ is W_{Ch} (see Remark 10.10); in addition, along curves $\mu_t = \mathsf{P}_t \rho \, \mathfrak{m}$ with $\rho \in L_+^\infty(X, \mathfrak{m})$ probability density, these distances are finite and the metric derivatives w.r.t. all these distances coincide a.e. in $(0, \infty)$ (see Corollary 7.5).

Building on this and refining the analysis made in [3], we prove in the subsequent Section 8 that the metric gradient flows in $\mathscr{P}^a(X)$ of the relative entropy functional

$$\mathrm{Ent}(\rho \, \mathfrak{m}) := \int \rho \log \rho \, \mathrm{d}\mathfrak{m} \tag{1.3}$$

w.r.t. the distances W_{d}, $W_{\mathrm{Ch},*}$ and W_{Ch} coincide with the Heat semigroup P in the P-invariant class of bounded probability densities under a mild lower semicontinuity assumption on $|\mathrm{D}^- \mathrm{Ent}|$, the descending slope of Ent w.r.t. W_{d}.

Section 9 contains a key stability result for Cheeger's energies, which deals with the case of a monotone family of $(\tau \times \tau)$-continuous distances d_i approximating from below d, as in the definition of extended metric-topological space. We prove convergence of the correspondent gradient flows and the formula

$$\mathrm{Ch} = \left(\inf_{i \in I} \mathrm{Ch}_i \right)_* \,,$$

where Ch_i is Cheeger's functional relative to d_i and $()_*$ denotes the lower semicontinuous envelope in $L^2(X, \mathfrak{m})$. As a byproduct, we get also convergence of the corresponding L^2 Heat flows. In view of the applications given in Section 12 we include in the convergence result also the case when d_i are semi distances, i.e. $\mathsf{d}_i(x, y) = 0$ does not imply $x = y$. This inclusion requires an adaptation of the construction of Ch to the semimetric setting.

In Section 10 we take the point of view of a strongly local and Markovian Dirichlet form $\mathcal{E}$ endowed with a carré du champs operator Γ. In this context the definitions of W_{Ch} and $W_{\mathrm{Ch},*}$ can be immediately adapted, formally replacing the minimal relaxed slope $|\mathrm{D}f|_w$ with $\sqrt{\Gamma(f)}$. Denoting by $W_{\mathcal{E}}$ and $W_{\mathcal{E},*}$ the corresponding extended distances, also in this context the K-gradient contractivity condition (1.1) yields that $W_{\mathcal{E}}$ is the length distance associated to $W_{\mathcal{E},*}$; furthermore, we prove in Theorem 10.14 that (1.1) implies K-contractivity of both squared distances w.r.t. to $\mathsf{P}^{\mathcal{E}}$; if $L^2(X, \mathfrak{m})$ is separable we prove a partial converse, namely K-contractivity of $W_{\mathcal{E}}^2$ implies the K-gradient contractivity (1.1).

Section 11 provides the EVI_K property (1.2) of $\mathsf{S} = \mathsf{P}^{\mathcal{E}}$ relative to the extended distances $W_{\mathcal{E},*}$ and $W_{\mathcal{E}}$. The proof first provides a duality estimate involving $W_{\mathcal{E},*}$ and then, using the self-improvement principle of Section 3 and the relation between $W_{\mathcal{E},*}$ and $W_{\mathcal{E}}$, the final result. In conjunction with the compactness properties of the sublevels of Ent, the EVI_K property yields geodesic convexity, relative to $W_{\mathcal{E}}$, of the sublevels of the entropy. In addition, when $K \geq 0$ also the sets $\{\rho\mathfrak{m} : \|\rho\|_\infty \leq c\}$ are convex and when $K > 0$ the Logarithmic Sobolev Inequality and the Talagrand Transport Inequality hold.

In Section 12 we analyze more in detail the connection between the metric and differentiable structures. More precisely, starting from an extended metric-topological space (X, d, τ) endowed with a reference measure $\mathfrak{m} \in \mathscr{P}(X)$ we can build Ch and, assuming Ch to be quadratic, ask whether Ch fits into the theory of Dirichlet forms; the answer is affirmative and we prove, following essentially [4], that Ch is a strongly local and Markovian Dirichlet form, and that $|\mathrm{D}f|_w^2$ corresponds to $\Gamma(f)$. Conversely, given $\mathcal{E}$ we can easily build an extended metric measure structure by selecting a family $\mathcal{A} \subset \{f : \Gamma(f) \leq 1\}$ of pointwise defined functions which separate

the points of X. It is important to understand to what extent these two constructions at the level of the energies and of the distances are each the inverse of the other, namely

$$\mathcal{E} \to \mathsf{d}_\mathcal{E} \to \mathrm{Ch}_{\mathsf{d}_\mathcal{E}} = \mathcal{E}? \qquad \mathsf{d} \to \mathrm{Ch} \to \mathsf{d}_{\mathrm{Ch}} = \mathsf{d}?$$

We know from [43] that, in general, even for Dirichlet forms and for length distances [42], we can't expect that the answer is always affirmative. In order to understand this question we identify special properties of Cheeger's energies Ch and of distances d_{Ch} associated to them, when one chooses as $\mathcal{A}$, in the construction of d_{Ch}, the class $\{f \in C_b(X) : \ |\mathrm{D}f|_w \leq 1\}$. For Cheeger's energies, the special property is the so-called τ-upper regularity, already identified in [5] and here proved and adapted to the extended setting: according to this property $|\mathrm{D}f|_w$ can be approximated by τ-upper semicontinuous functions bounding the gradients of approximating functions f_n relative to finite distances, see Definition 12.4 for the precise statement. At the level of distances, the special property we need is that functions in the class $\{f \in C_b(X) : \ |\mathrm{D}f|_w \leq 1\}$ have to be 1-Lipschitz w.r.t. d.

We prove that $\mathcal{E} \leq \mathrm{Ch}_{\mathsf{d}_\mathcal{E}}$ and that equality holds iff the choice of $\mathcal{A}$ ensures the τ-upper regularity of $\mathcal{E}$ and that when this happens several other identifications occur, see Theorem 12.5 for the precise statement; the proof, in part adapted from [5], relies on the identification results established in the previous sections and particularly on Section 9. Notice that in the "regular" setting of [5] where τ coincides with the topology induced by $\mathsf{d}_\mathcal{E}$, τ-regularity can be also obtained as a consequence of the Bakry-Émery $\mathsf{BE}(K, \infty)$ condition and a weak Feller property of P.

On the other hand, in Theorem 12.8 we prove that d_{Ch} is always larger than d, and that equality holds if and only any function in $\{f \in C_b(X) : \ |\mathrm{D}f|_w \leq 1\}$ is 1-Lipschitz w.r.t. d. These results are independent of the doubling and Poincaré assumption considered in [30], [31] and generalize those of [5] to the extended setup.

Section 13 describes classical examples of extended spaces (Wiener spaces, configuration spaces) and shows how they fit in our framework.

Let us conclude by pointing out potential developments that we plan to investigate, at least in part, in [8]. The first one deals with the so-called measurable distances, namely distances which are pointwise defined only at the level of subsets of positive $\mathfrak{m}$-measure of the space. This class of distances appears for instance in [28], in connection with the short time behaviour of the heat kernel, and it has been deeply investigated in [48], [49], in particular looking for (extended) metric realizations of measurable distances. Another direction comes from Gromov-Milman's theory of concentration, nicely developed in the recent monograph [41]: indeed, the notion of pyramid of metric measure spaces investigated in [41] displays some analogy with the monotone approximation property of the extended distance used in our axiomatization. In addition, the convergence result proved in Section 9 at the level of the Cheeger energies (for the special case of the monotone approximation) should be compared with the various convergence results (which essentially use, instead, the algebra of bounded 1-Lipschitz functions) developed in [41].

2. Preliminaries

2.1. **Measure-theoretic notation, Radon measures, weak and narrow topology.** For a Hausdorff topological space (X, τ), we denote by $C_b(X)$ the space of bounded continuous functions $f : (X, \tau) \to \mathbb{R}$ and by $\mathscr{B}(\tau)$ the Borel σ-algebra of τ. Throughout this paper

$$\mathscr{P}(X) \text{ denotes the class of Radon probability measures in } X, \qquad (2.1)$$

i.e. Borel probability measures μ having the property that

for any $B \in \mathscr{B}(\tau)$ and any $\epsilon > 0$ there exists a compact set $K \subset B$ with $\mu(B \setminus K) < \epsilon$ $\quad$ (2.2)

Notice that the Radon property is in general stronger than the tightness one, for which the inner approximation (2.2) is required only for $B = X$.

Radon measures have stronger additivity and continuity properties in connection with open sets and lower semicontinuous functions; in particular we shall use this version of the monotone convergence theorem (see [14, Lem. 7.2.6])

$$\lim_{i \in I} \int f_i \, \mathrm{d}\mu = \int \lim_{i \in I} f_i \, \mathrm{d}\mu \tag{2.3}$$

valid for Radon measures μ and for nondecreasing nets of τ-lower semicontinuous and equi-bounded functions f_i. By truncation, we can apply the same property to nondecreasing nets of τ-lower semicontinuous $f_i : X \to [0, \infty]$.

By the very definition of Radon topological space [39, Ch. II, Sect. 3], every Borel measure in a Radon space is Radon: such class of spaces includes locally compact spaces with a countable base of open sets, Polish, Lusin and Souslin spaces [39, Thm. 9 & 10, p. 122]. In particular the notation (2.1) is consistent with the standard one adopted e.g. in [3, 2, 45], where Polish or second countable locally compact spaces are considered.

The narrow topology on $\mathscr{P}(X)$ can be defined as the coarsest topology for which all maps

$$\mu \mapsto \int h \, \mathrm{d}\mu \qquad \text{from } \mathscr{P}(X) \text{ into } \mathbb{R} \tag{2.4}$$

are lower semicontinuous as $h : X \to \mathbb{R}$ varies in the set of bounded lower semicontinuous functions [39, p. 370]. It can be shown [39, p. 371] that it is a Hausdorff topology on $\mathscr{P}(X)$; when (X, τ) is completely regular, i.e.

$$\text{for any closed set } F \subset X \text{ and any } x_0 \in X \setminus F$$
$$\text{there exists } f \in C_b(X) \text{ with } f(x_0) > 0 \text{ and } f \equiv 0 \text{ on } F, \tag{2.5}$$

the narrow topology coincides with the usual weak one, induced by the duality with $C_b(X)$. In fact, in a completely regular space every bounded lower semicontinuous function h is the upper envelope of the directed set $D_h := \{f \in C_b(X) : f \le h\}$, so that (2.3) shows that $\int h \, \mathrm{d}\mu = \sup \{ \int f \, \mathrm{d}\mu : f \in D_h \}$.

One of the advantages to use the narrow topology in $\mathscr{P}(X)$ when (X, τ) is a Hausdorff topological space is the following sufficient condition for the compactness [39, Theorem 3, p. 379] (in completely regular spaces it is a consequence of Prokhorov theorem).

Theorem 2.1. *Let (X, τ) be a Hausdorff topological space. Assume that a collection $\mathcal{M} \subset \mathscr{P}(X)$ is equi-tight, i.e.*

$$\text{for every } \varepsilon > 0 \text{ there exists a compact set } K_\varepsilon \subset X \text{ such that } \quad \sup_{\mu \in \mathcal{M}} \mu(X \setminus K_\varepsilon) \le \varepsilon . \tag{2.6}$$

Then $\mathcal{M}$ has limit points in the class $\mathscr{P}(X)$ w.r.t. the narrow topology (in particular, the weak topology induced by $C_b(X)$ when (X, τ) is completely regular).

2.2. **Transport plans, gluing, optimal transport and duality.** Let (X, τ) and (Y, σ) be Hausdorff topological spaces and let $\mu \in \mathscr{P}(X)$. We say that a map $f : X \to Y$ is Lusin μ-measurable (see e.g. [39, Chap. I, Sect. 5]) if

$$\text{for every compact set } K \subset X \text{ and every } \delta > 0 \text{ there exists a compact set}$$
$$K_\delta \subset K \text{ such that } f \text{ restricted to } K_\delta \text{ is continuous and } \mu(K \setminus K_\delta) \le \delta . \tag{2.7}$$

Notice that since μ is a Radon measure the approximation property (2.7) holds in fact for every $K \in \mathscr{B}(\tau)$.

If f is Lusin μ-measurable than it is also Borel μ-measurable (i.e. $f^{-1}(B)$ is μ-measurable for every $B \in \mathscr{B}(\sigma)$); the converse is known to be true if (Y, σ) is separable and metrizable [39, Thm. 5, p. 26].

If f is Lusin μ-measurable we denote by $f_\sharp \mu \in \mathscr{P}(Y)$ the push-forward of μ via f: it is the Radon measure defined by $f_\sharp \mu(B) := \mu(f^{-1}(B))$ for every Borel set $B \in \mathscr{B}(Y)$ (Lusin's μ-measurability of f is assumed in order to guarantee the Radon property of $f_\sharp \mu$).

If $\mu, \nu \in \mathscr{P}(X)$, we will denote by $\Gamma(\mu, \nu)$ the class of admissible transport plans between μ and ν, i.e. Radon probability measures in $\mathscr{P}(X \times X)$ with marginals μ and ν respectively:

$$
\Gamma(\mu, \nu) = \Big\{ \boldsymbol{\pi} \in \mathscr{P}(X \times X) : \ \pi(A \times X) = \mu(A) \ \text{ for all } A \in \mathscr{B}(\tau)
$$
$$
\pi(X \times B) = \nu(B) \ \text{ for all } B \in \mathscr{B}(\tau) \Big\} . \tag{2.8}
$$

It is worth noticing that $\Gamma(\mu, \nu)$ is non empty (since it contains the unique Radon extension to $\mathscr{B}(X \times X)$ of the product measure $\mu \times \nu$, see [39, p. 73]), convex and compact with respect to the narrow topology, by Theorem 2.1.

We shall use the following gluing lemma:

Lemma 2.2 (Gluing lemma). *Let $I = \{0, 1, \ldots, N\}$, $N \geq 2$, and let $(X_i)_{i \in I}$ be Hausdorff topological spaces with $\boldsymbol{X} := \Pi_{i \in I} X_i$ and corresponding projection p^i. Let $\boldsymbol{\pi}_i$, $i = 1, \ldots, N$, be Radon measures in $X_{i-1} \times X_i$ satisfying the compatibility conditions*

$$
\int \phi(y) \, \mathrm{d}\boldsymbol{\pi}_i(x, y) = \int \phi(x) \, \mathrm{d}\boldsymbol{\pi}_{i+1}(x, y) \qquad \forall \phi \in C_b(X_i) \qquad \text{for every } i \in \{1, \ldots, N-1\} .
$$

Then there exist Radon measures $\boldsymbol{\pi}$ in $\boldsymbol{X}$ such that $(\mathsf{p}^{i-1}, \mathsf{p}^i)_\sharp \boldsymbol{\pi} = \boldsymbol{\pi}_i$ for every $i \in \{1, \ldots, N-1\}$. The same property (with obvious modifications) holds in the case $I = \mathbb{N}$ of a countable set of indexes.

Proof. When I is finite, the proof is well-known in Polish spaces (see for instance [2, Lem. 5.3.2]) via disintegrations; however an alternative proof via Hahn-Banach theorem and Riesz-Markov-Kakutani theorem [14, Thm. 7.3.10 and Thm. 7.10.4] is possible in compact Hausdorff spaces, as indicated in [45, Exer. 7.9] in the case $N = 2$, and then proceeding by induction. By a simple exhaustion argument the result extends to Radon measures in Hausdorff topological spaces.

In the case $I = \mathbb{N}$ we argue as in the proof of [2, Lem. 5.3.4], applying the general version of Kolmogorov-Prokhorov theorem given in [39, Thm. 21, p. 74 and its Corollary p. 81]. $\square$

In the class of Radon probability measures in X, we consider the optimal transport problem

$$
\inf \Big\{ \int_{X \times X} \mathsf{c} \, \mathrm{d}\boldsymbol{\pi} : \ \boldsymbol{\pi} \in \Gamma(\mu, \nu) \Big\} \tag{2.9}
$$

where c is a bounded cost function defined in $X \times X$. In the following proposition we denote by $\int_* f \, \mathrm{d}\nu$ the inner integral, namely the supremum of $\int g \, \mathrm{d}\nu$ among all ν-measurable g with $g \leq f$ pointwise. Even though the natural setting for duality theorems is provided by costs measurable w.r.t. the product σ-algebra, we will need to apply the duality theorem with lower semicontinuous costs $\mathsf{c} \geq 0$. In this case the duality theorem still holds, when Radon measures are involved; this can be seen, for instance, proving via (2.3) the existence of $\widetilde{\mathsf{c}} \leq \mathsf{c}$ measurable w.r.t. to $\mathscr{B}(\tau) \times \mathscr{B}(\tau)$ with $\int \mathsf{c} \, \mathrm{d}\boldsymbol{\pi} = \int \widetilde{\mathsf{c}} \, \mathrm{d}\boldsymbol{\pi}$.

Proposition 2.3 (Duality). *Let $\mathsf{c} : X \times X \to \mathbb{R}$ be a bounded and either $(\tau \times \tau)$-lower semicontinuous or $\mathscr{B}(\tau) \times \mathscr{B}(\tau)$-measurable function. For all $\mu, \nu \in \mathscr{P}(X)$ one has*

$$
\inf \Big\{ \int_{X \times X} \mathsf{c} \, \mathrm{d}\boldsymbol{\pi} : \ \boldsymbol{\pi} \in \Gamma(\mu, \nu) \Big\} = \sup \Big\{ \int_* \psi \, \mathrm{d}\nu - \int \phi \, \mathrm{d}\mu \Big\} \tag{2.10}
$$

where the supremum runs in any of the following three classes:

(a) *ϕ, ψ $\mathscr{B}(\tau)$-measurable with $\psi(y) - \phi(x) \leq \mathsf{c}(x, y)$ and $\int |\psi| \, \mathrm{d}\nu + \int |\phi| \, \mathrm{d}\mu < \infty$;*
(b) *ϕ, ψ bounded $\mathscr{B}(\tau)$-measurable with $\psi(y) - \phi(x) \leq \mathsf{c}(x, y)$;*
(c) *(ϕ, ψ) with $\psi(y) = \phi^{\mathsf{c}}(y) = \inf_x \phi(x) + \mathsf{c}(x, y)$ and ϕ belonging to the class*

$$
\mathcal{F} := \Big\{ \phi : X \to [0, \infty) : \ \phi \in C(K) \text{ for some compact } K \subset X, \ \phi|_{X \setminus K} \equiv c \geq \max_K \phi \Big\} . \tag{2.11}
$$

Notice that in the cases (a) and (b) one can replace the inner integral $\int_* \psi \, \mathrm{d}\nu$ with $\int \psi \, \mathrm{d}\nu$ in (2.10).

Proof. At this level of generality, the validity of (2.10) with the choice (a) has been proved in [29, Prop. 1.31, Thm. 2.14] for both classes of costs c. (b) still follows by [29, Lem. 1.8]: in fact one can assume that ϕ, ψ take values in the interval $[\inf \mathsf{c} - \frac{1}{2} \sup \mathsf{c}, \frac{1}{2} \sup \mathsf{c}]$. Let us pass now from (b) to (c). To this aim we can use once more the Radon assumption on μ, which implies the validity of Lusin's theorem (see [14, Thm. 7.1.13]) to find $\phi_n \in \mathcal{F}$ equibounded with $\mu(\{\phi \neq \phi_n\}) \downarrow 0$ and $\phi_n \geq \phi$ pointwise. Since

$$\int_* \phi_n^{\mathsf{c}} \, \mathrm{d}\nu \geq \int_* \phi^{\mathsf{c}} \, \mathrm{d}\nu \geq \int \psi \, \mathrm{d}\nu$$

we conclude.   □

In some sections of this paper a reference measure $\mathfrak{m} \in \mathscr{P}(X)$ will be fixed; we shall denote by $\mathscr{P}^a(X)$ the subclass of measures $\mu \ll \mathfrak{m}$ and by Ent the relative entropy w.r.t. $\mathfrak{m}$, equal to $+\infty$ on $\mathscr{P}(X) \setminus \mathscr{P}^a(X)$ and given by

$$\mathrm{Ent}(\rho\mathfrak{m}) := \int \rho \log \rho \, \mathrm{d}\mathfrak{m} \tag{2.12}$$

otherwise. Since $\mathfrak{m}$ will be fixed, we do not emphasize it in the notations $\mathscr{P}^a(X)$, Ent and we will also use the short notation $\|f\|_p$ for $\|f\|_{L^p(X,\mathfrak{m})}$, $p \in [1,\infty]$.

3. EXTENDED METRIC SPACES

In this section we introduce some basic analytic tools for analysis in metric spaces. Most of them extend in a natural way to extended metric spaces, defined below.

Definition 3.1 (Extended metric spaces). *We say that (X, d) is an extended metric space if $\mathsf{d} : X \times X \to [0, \infty]$ is a symmetric function satisfying the triangle inequality, with $\mathsf{d}(x, y) = 0$ iff $x = y$.*

When the condition $\mathsf{d}(x, y) = 0$ iff $x = y$ is weakened to $\mathsf{d}(x, x) = 0$ we say that (X, d) is a (extended) semimetric space.

The main example of extended semimetric space we have in mind arises from the construction

$$\mathsf{d}(x, y) := \sup_{f \in \mathcal{A}} |f(x) - f(y)| \,, \tag{3.1}$$

where $\mathcal{A}$ is any class of bounded functions $f : X \to \mathbb{R}$; it is an extended metric space iff $\mathcal{A}$ separates the points of X. Other natural examples arise by the construction of length distances (derived from distances or from action minimization), which need not to be finite.

Since an extended metric space can be seen as the disjoint union of the equivalence classes induced by the equivalence relation

$$x \sim y \qquad \Longleftrightarrow \qquad \mathsf{d}(x, y) < \infty$$

and since the equivalence classes are indeed metric spaces, many results and definitions extend with no effort to extended metric spaces. For instance, we say that an extended metric space (X, d) is complete (resp. geodesic, length,...) if all metric spaces $X_{[x]} = \{y : \ y \sim x\}$ are complete (resp. geodesic, length,...). In particular any extended metric space has a unique, up to isometries, (extended) metric completion.

3.1. Absolutely continuous curves and upper gradients.
For $D \subset \mathbb{R}$ we denote by X^D the space of maps $\eta : D \to X$ (no continuity is required in general) and by $\mathrm{e}_t : X^D \to X$ the evaluation maps at time $t \in D$, namely $\mathrm{e}_t(\eta) := \eta(t)$.

Let $UC(D; (X, \mathsf{d})) \subset X^D$ be the space of d-uniformly continuous maps. For $p \in (1, \infty)$, we denote $AC^p(D; (X, \mathsf{d})) \subset UC(D; (X, \mathsf{d}))$ the subspace of p-absolutely continuous maps w.r.t. d, satisfying

$$\mathscr{A}_p(\eta, D) := \sup \left\{ \sum_{i=0}^{n-1} \frac{\mathsf{d}^p(\eta(t_{i+1}), \eta(t_i))}{(t_{i+1} - t_i)^{p-1}} : \ t_0 < t_1 < \cdots < t_{n-1} < t_n, \ t_i \in D \right\} < \infty \,. \tag{3.2}$$

Notice that, thanks to the triangle inequality, the supremum above can also be realized as a monotone nondecreasing limit in the directed set of partitions of D, with the order induced

by the set-theoretic inclusion. When the domain D is clear from the context we simplify the notation, writing $\mathscr{A}_p(\eta)$. If (X, d) is complete, it is easily seen that any $\eta \in UC(D; (X, \mathsf{d}))$ has a unique d-continuous extension to $\overline{D}$, and this extension is still d-uniformly continuous; if $\eta \in AC^p(D; (X, \mathsf{d}))$ then the extension belongs to $AC^p(\overline{D}; (X, \mathsf{d}))$ and the p-action $\mathscr{A}_p$ remains the same. The metric derivative $|\dot{\eta}| : (0, T) \to [0, \infty]$ of $\eta \in AC^p([0, T]; (X, \mathsf{d}))$ is the Borel map defined by

$$|\dot{\eta}|(t) := \limsup_{s \to t} \frac{\mathsf{d}(\eta(s), \eta(t))}{|s - t|}$$

and it can be proved (see for instance [2, Thm. 1.1.2]) that the $\limsup$ is a limit a.e. in $(0, T)$, that $|\dot{\eta}| \in L^p(0, T)$ and that $\mathsf{d}(\eta(s), \eta(t)) \le \int_s^t |\dot{\eta}|(r) \, \mathrm{d}r$ for all $0 \le s \le t \le T$. Furthermore, $|\dot{\eta}|$ is the smallest L^1 function with this property, up to Lebesgue negligible sets, and one can easily prove that

$$\mathscr{A}_p(\eta) = \int_0^T |\dot{\eta}(t)|^p \, \mathrm{d}t \qquad \forall \eta \in AC^p([0, T]; (X, \mathsf{d})) \ . \tag{3.3}$$

We also recall the notion of upper gradient: we say that $g : X \to [0, \infty]$ is an upper gradient of $f : X \to \mathbb{R}$ (relative to d) if $t \mapsto g(\eta(t))|\dot{\eta}|(t)$ is Lebesgue measurable in $(0, 1)$ and

$$|f(\eta(1)) - f(\eta(0))| \le \int_0^1 g(\eta(r))|\dot{\eta}|(r) \, \mathrm{d}r$$

for any $\eta \in AC^p([0, 1]; (X, \mathsf{d}))$, $p \in (1, \infty)$ (the dependence on p is harmless, thanks to reparameterizations).

3.2. Gradient flows.

Now we introduce the main concepts of gradient flows used in this paper: the first one is based on the energy-dissipation inequality and the second is characterized by a family of evolution variational inequalities, see [2] and [18] for much more on this topic. Both can be easily generalized to the extended setting and the case of gradient flows in Hilbert spaces, detailed in Proposition 3.4, is a very particular and important example.

Let $F : X \to \mathbb{R} \cup \{+\infty\}$ with (non empty) domain $D(F) = \{F < \infty\}$. The slope $|\mathrm{D}F|$ and the descending slope $|\mathrm{D}^- F|(x)$ of F at $x \in D(F)$ are respectively defined by

$$|\mathrm{D}F|(x) := \limsup_{y \to x} \frac{|F(y) - F(x)|}{\mathsf{d}(y, x)} \ , \qquad |\mathrm{D}^- F|(x) := \limsup_{y \to x} \frac{[F(y) - F(x)]^-}{\mathsf{d}(y, x)} \ , \tag{3.4}$$

with the convention $|\mathrm{D}F|(x) = |\mathrm{D}^- F|(x) = 0$ if x is a d-isolated point.

We say that a locally absolutely continuous curve $x \in AC^2_{\mathrm{loc}}((0, \infty); (D(F), \mathsf{d}))$ is a metric gradient curve of F in the energy-dissipation sense if the *Energy Dissipation Inequality*

$$F(x(t)) + \frac{1}{2} \int_s^t \left(|\dot{x}|^2(r) + |\mathrm{D}^- F|^2(x(r)) \right) \, \mathrm{d}r \le F(x(s)) \tag{EDI}$$

holds for all $s, t \in (0, \infty)$ with $s < t$ and also for $s = 0$ when the curve x admits a continuous extension (still denoted by x) to $[0, \infty)$ (recall the definition (3.4) of $|\mathrm{D}^- F|$).

Notice that if $|\mathrm{D}^- F|$ is an upper gradient of F we can use the inequality

$$F(x(s)) \le F(x(t)) + \int_s^t |\mathrm{D}^- F|(x(r))|\dot{x}(r)| \, \mathrm{d}r$$

to obtain that equality holds in (EDI), that $F \circ x$ is locally absolutely continuous in $(0, \infty)$ and that $|\dot{x}|^2 = |\mathrm{D}^- F|^2 \circ x = -(F \circ x)'$ a.e. in $(0, \infty)$. When x has a continuous extension to $[0, \infty)$ and $F(x_0) < \infty$, the absolute continuity holds in all compact intervals of $[0, \infty)$.

For $K \in \mathbb{R}$ we introduce the function $\mathrm{I}_K : [0, \infty) \to [0, \infty)$ defined by

$$\mathrm{I}_0(t) = t, \qquad \mathrm{I}_K(t) := \int_0^t \mathrm{e}^{Ks} \, \mathrm{d}s = \frac{\mathrm{e}^{Kt} - 1}{K} \qquad \text{if } K \ne 0 \ .$$

In the following definition of EVI_K gradient flow we consider some lower semicontinuity conditions that are automatically implied by the local absolute continuity of x when F is lower semicontinuous and d is a finite distance. Moreover, the specification of the initial condition is

made in such a way that even initial conditions not in $\overline{D(F)}$, but at a finite distance from $D(F)$, can be considered.

Definition 3.2 (EVI_K gradient flows). *Let $K \in \mathbb{R}$. We say that $x \in AC_{\mathrm{loc}}^2((0, \infty); (D(F), \mathsf{d}))$ is a EVI_K gradient curve of F if $t \mapsto F(x_t)$ is lower semicontinuous in $(0, \infty)$ and for all $y \in D(F)$ satisfying $\mathsf{d}(y, x_t) < \infty$ for some (and then all) $t \in (0, \infty)$ one has*

$$\frac{\mathrm{d}^+}{\mathrm{d}t} \frac{1}{2} \mathsf{d}^2(x_t, y) + \frac{K}{2} \mathsf{d}^2(x_t, y) \leq F(y) - F(x_t) \qquad \forall t > 0 \,, \qquad \text{(EVI}_K)$$

where $\mathrm{d}^+/\mathrm{d}t$ denotes the upper right derivative.
If $\bar{x}$ has finite distance from $D(F)$, we say that x starts from $\bar{x}$ if $\liminf_{t \downarrow 0} F(x_t) \geq F(\bar{x})$ and $\lim_{t \downarrow 0} \mathsf{d}(x_t, y) \to \mathsf{d}(\bar{x}, y)$ for every $y \in \overline{D(F)}$.

Let us now point out some direct consequences of the definition of EVI_K gradient curve.
Monotonicity of the energy, uniqueness and contractivity. It is not difficult to show that, for EVI_K gradient curves x, the map $t \mapsto F(x_t)$ is non-increasing in $(0, \infty)$ and $F(x_t) \to F(\bar{x})$ if x starts from $\bar{x}$. Moreover one has the contractivity property

$$\mathsf{d}(x_{t+s}, y_{t'+s}) \leq e^{-Ks} \mathsf{d}(x_t, y_{t'}) \qquad s, \, t, \, t' > 0 \,.$$

By approximation this inequality can be extended to the case when either $t = 0$ or $t' = 0$ (but not both), for EVI_K curves starting respectively from $\bar{x}$, $\bar{y}$.
If $t = t' = 0$ the inequality holds in the weaker form

$$\mathsf{d}(x_s, y_s) \leq e^{-Ks} \inf_{y \in D(F)} \mathsf{d}(\bar{x}, y) + \mathsf{d}(y, \bar{y}) \qquad s \geq 0 \,,$$

which reduces to the standard one when at least one of the initial points belongs to $\overline{D(F)}$. In particular the EVI_K gradient curve starting from $\bar{x} \in \overline{D(F)}$ is unique and satisfies $\mathsf{d}(x(t), \bar{x}) \to 0$ as $t \downarrow 0$.
Integral version and regularization. Integrating in $(0, t)$ the differential inequality (EVI$_K$) written in the form $\frac{\mathrm{d}^+}{\mathrm{d}t} \left(\frac{e^{Kt}}{2} \mathsf{d}^2(x_t, y) \right) \leq e^{Kt}(F(y) - F(\mathsf{S}_t x))$ and using the monotonicity of $F(\mathsf{S}_t x)$ and the convergence $\mathsf{d}(x_s, y) \to \mathsf{d}(\bar{x}, y)$ as $s \downarrow 0$ we get

$$\frac{1}{2} \mathsf{d}^2(x_t, y) - \frac{e^{-Kt}}{2} \mathsf{d}^2(\bar{x}, y) \leq \mathrm{I}_{-K}(t)\big(F(y) - F(x_t)\big) \qquad (3.5)$$

for every $y \in D(F)$ with $\mathsf{d}(\bar{x}, y) < \infty$, $t > 0$.
In particular we obtain the regularization estimate (see also (3.7) below)

$$F(x_t) \leq F(y) + \frac{1}{2\mathrm{I}_K(t)} \mathsf{d}^2(\bar{x}, y) \qquad \forall y \in D(F), \, t > 0 \,. \qquad (3.6)$$

Definition 3.3 (EVI_K-semigroup of F). *Let $D \subset X$ and let $\mathsf{S} : D \times [0, \infty) \to D$ be a semigroup. We say that S is an EVI_K gradient flow of F in D if:*

  (i) *$D \supset \overline{D(F)}$ and every $x \in D$ has finite distance from $D(F)$;*
  (ii) *for every $x \in D$ one has that $t \mapsto \mathsf{S}_t x$ is an EVI_K gradient curve of F starting from x.*
  (iii) *S is K-contractive in D, i.e. $\mathsf{d}(\mathsf{S}_t x, \mathsf{S}_t y) \leq e^{-Kt} \mathsf{d}(x, y)$ for every $x, y \in D$.*

It is not hard to show that if (i) and (ii) hold, then condition (iii) is always satisfied in $\overline{D(F)}$, so that it is usually omitted in all the cases when $D = \overline{D(F)}$, in particular when $D(F)$ is dense in X.

In general existence is much harder to prove and depends on structural properties of (X, d) and F. A particularly important case is provided by lower semicontinuous convex functionals in Hilbert spaces, that we are now briefly recalling (see for instance [2, Sec. 1.4] for the "metric" approach and [16] for the classic formulation via maximal monotone operators).

If $(X, \| \cdot \|)$ is a Hilbert space, recall that the subdifferential $\partial F(x)$ of a convex functional $F : X \to (-\infty, \infty]$ at $x \in D(F)$ is the convex closed set (possibly empty)

$$\partial F(x) := \{\xi \in X : \ F(y) \geq F(x) + \langle \xi, y - x \rangle \ \forall y \in X\} \,.$$

Proposition 3.4 (Gradient flows of convex functionals in Hilbert spaces)**.** *Assume that X is a Hilbert space, with distance d induced by the norm, and that $F : X \to [0, \infty]$ is convex and lower semicontinuous. Then, the following properties hold:*

(a) *the concepts of metric gradient curve and EVI_0 gradient flow starting from $\bar{x}$ coincide, for locally absolutely continuous curves $x : [0, \infty) \to X$ with $x_0 = \bar{x} \in \overline{D(F)}$;*

(b) *For all $\bar{x} \in \overline{D(F)}$, there exists a unique metric gradient curve $x(t)$ of F starting from $\bar{x}$. The map $\mathsf{S}_t : \bar{x} \mapsto x(t)$ defines an EVI_0 semigroup of F in $\overline{D(F)}$.*

(c) *If $x(t)$ is a metric gradient curve of F, then $-x'_+(t) = \lim_{h \downarrow 0}(x(t) - x(t+h))/h$ exists for all $t > 0$ and it coincides with the element with minimal norm of $\partial F(x(t))$.*

(d) *If $x(t)$, $y(t)$ are metric gradient curves of F, then $\mathsf{d}(x(t), y(t)) \leq \mathsf{d}(x(s), y(s))$, $0 < s \leq t < \infty$.*

(e) *If $x(t)$ is the metric gradient curve starting from $\bar{x}$ the following regularization estimates hold:*

$$F(x(t)) \leq F(z) + \frac{1}{2t}\mathsf{d}^2(z, \bar{x}) \,, \quad |\mathrm{D}^- F|^2(x(t)) \leq |\mathrm{D}^- F|^2(\bar{x}) + \frac{1}{t^2}\mathsf{d}^2(z, \bar{x}) \tag{3.7}$$

for all $z \in D(F)$ and $t > 0$.

3.3. EVI flows, length distances and geodesic convexity.

If (X, d) is an extended metric space, we denote by $\mathsf{d}_\ell \geq \mathsf{d}$ the extended length distance associated to d: since we are not making any completeness assumption at this level, it can be defined as

$$\mathsf{d}_\ell(y, z) := \sup_{\varepsilon > 0} \mathsf{d}^\varepsilon(y, z) = \lim_{\varepsilon \downarrow 0} \mathsf{d}^\varepsilon(y, z), \quad \text{where}$$

$$\mathsf{d}^\varepsilon(y, z) := \inf \Big\{ \sum_{n=1}^{N} \mathsf{d}(x_n, x_{n-1}) : x_0 = y, \ x_N = z, \ \mathsf{d}(x_{n-1}, x_n) \leq \varepsilon \Big\} . \tag{3.8}$$

A second way to generate a length distance from d consists in minimizing the length of all the absolutely continuous curves connecting two points: we set

$$\bar{\mathsf{d}}_\ell(y, z) := \inf \Big\{ \int_0^1 |\dot{x}|(t)\, \mathrm{d}t : \ x \in AC([0, 1]; (X, \mathsf{d})), \ x_0 = y, \ x_1 = z \Big\} . \tag{3.9}$$

Notice that if $x \in AC([0, 1]; (X, \mathsf{d}))$ is an absolutely continuous curve connecting y to z we easily get

$$\mathsf{d}_\ell(y, z) \leq \int_0^1 |\dot{x}|(t)\, \mathrm{d}t \,, \quad \text{so that } \mathsf{d}_\ell \leq \bar{\mathsf{d}}_\ell . \tag{3.10}$$

Motivated by (3.10) we call $\bar{\mathsf{d}}_\ell$ the *upper* length distance associated to d.

(X, d) is called a *length* space if $\mathsf{d}_\ell = \mathsf{d}$. It is not difficult to check that d is a length distance if and only if for every $x, y \in X$ with $\mathsf{d}(x, y) < \infty$ and all $\varepsilon > 0$ there exists an ε-middle point z_ε, characterized by

$$\mathsf{d}(x, z_\varepsilon) \leq \frac{1}{2}\mathsf{d}(y, z) + \varepsilon \,, \quad \mathsf{d}(z_\varepsilon, y) \leq \frac{1}{2}\mathsf{d}(y, z) + \varepsilon . \tag{3.11}$$

With this in mind, one can easily check that d_ℓ and $\bar{\mathsf{d}}_\ell$ are length distances. If moreover (X, d) is complete one has $\mathsf{d}_\ell = \bar{\mathsf{d}}_\ell$ (a simple proof can be achieved by selecting, given x and y, $\varepsilon/4^i$-midpoints, $i \geq 0$, recursively; thanks to the completeness, the process converges to an absolutely continuous curve from x to y with length less than $\mathsf{d}_\ell(x, y) + 2\varepsilon$). In the next theorem we describe a new self-improvement principle for EVI_K gradient flows.

Theorem 3.5 (Self-improvement of EVI_K)**.** *Let $F : X \to \mathbb{R} \cup \{+\infty\}$ be a proper functional and let S be an EVI_K gradient flow of F in X relative to d. Then S is an EVI_K gradient flow of F relative to d_ℓ and $\bar{\mathsf{d}}_\ell$.*

Proof. We discuss the case of d_ℓ; the proof for $\bar{\mathsf{d}}_\ell$ is analogous, working with continuous rather than discrete curves.

Let us fix $y \in D(F)$, $x \in X$ with $\mathsf{d}_\ell(x, y) < \infty$; the existence of ε-middle points as in (3.11) for arbitrary couples at finite d_ℓ-distance easily shows that for every $\varepsilon > 0$ and $N \geq 1$ there exist points x_i, $i = 0, \ldots, N$, with $x_0 = y$, $x_N = x$ and $\mathsf{d}(x_i, x_{i+1}) \leq \mathsf{d}_\ell(x, y) 2^\varepsilon/N$, $i = 0, \ldots, (N-1)$.

Set $s_i := i/N$, $x_i^t := \mathsf{S}_{ts_i} x_i$, and $\delta = 1/N$; notice that for $t > 0$ we have $x_n^t \in D(F)$, by our assumptions on S. Using first the contractivity of S and then (3.5) (with $y = x_i^t = \mathsf{S}_{ts_i} x_i$) we obtain

$$
\frac{1}{2}\mathsf{d}^2(x_{i+1}^t, x_i^t) - \mathrm{e}^{-Kt\delta}\mathrm{e}^{-2Kts_i}\frac{1}{2}\mathsf{d}^2(x_{i+1}, x_i)
$$
$$
\leq \frac{1}{2}\mathsf{d}^2(\mathsf{S}_{t\delta}\mathsf{S}_{ts_i}x_{i+1}, \mathsf{S}_{ts_i}x_i) - \mathrm{e}^{-Kt\delta}\frac{1}{2}\mathsf{d}^2(\mathsf{S}_{ts_i}x_{i+1}, \mathsf{S}_{ts_i}x_i)
$$
$$
\leq \mathrm{I}_{-K}(t\delta)\big[F(x_i^t) - F(x_{i+1}^t)\big] . \tag{3.12}
$$

Since $\delta^{-1} = N$ we have

$$
\sum_{i=0}^{N-1} \frac{1}{\delta}\mathsf{d}^2(x_{i+1}^t, x_i^t) \geq \left(\sum_{i=0}^{N-1} \mathsf{d}(x_{i+1}^t, x_i^t)\right)^2 \tag{3.13}
$$

$$
\sum_{i=0}^{N-1} \frac{1}{\delta}\mathsf{d}^2(x_{i+1}^t, x_i^t) \geq \max_i \frac{\mathsf{d}^2(x_{i+1}^t, x_i^t)}{\delta} \tag{3.14}
$$

and

$$
\sum_{i=0}^{N-1} \frac{1}{\delta}\mathrm{e}^{-Kt\delta}\mathrm{e}^{-2Kts_i}\mathsf{d}^2(x_{i+1}, x_i) \leq 2^{2\varepsilon}\mathrm{e}^{-Kt\delta}\mathsf{d}_\ell^2(x, y)\sum_i \delta\mathrm{e}^{-2Kts_i}
$$
$$
\leq 2^{2\varepsilon}\mathrm{e}^{-Kt\delta}\mathsf{d}_\ell^2(x, y)\Big(\int_0^1 \mathrm{e}^{-2Kts}\,\mathrm{d}s + \omega(\delta)\Big)
$$

with $\omega(\delta) \to 0$ as $\delta \to 0$. Thus, dividing by δ and summing up (3.12) we obtain

$$
\frac{1}{2}\sum_{i=0}^{N-1} \frac{1}{\delta}\mathsf{d}^2(x_{i+1}^t, x_i^t) \leq 2^{2\varepsilon}\mathrm{e}^{-Kt\delta}\frac{1}{2}\mathsf{d}_\ell^2(x, y)\Big(t^{-1}\mathrm{I}_{-2K}(t) + \omega(\delta)\Big) + \frac{\mathrm{I}_{-K}(\delta t)}{\delta}\big(F(y) - F(\mathsf{S}_t x)\big) .
$$

Taking the limit along a family of $\varepsilon \downarrow 0$ and $N \uparrow \infty$, (3.13) and (3.14) show that the $\liminf$ of the left hand side of the last inequality provides un upper bound of $\frac{1}{2}\mathsf{d}_\ell^2(\mathsf{S}_t x, y)$, and this yields

$$
\frac{1}{2}\mathsf{d}_\ell^2(\mathsf{S}_t x, y) - \frac{\mathrm{I}_{-2K}(t)}{2t}\mathsf{d}_\ell^2(x, y) \leq t(F(y) - F(\mathsf{S}_t x)) .
$$

This implies the differential inequality at $t = 0$ and then (EVI_K) for the extended distance d_ℓ, thanks to the semigroup property of S. $\qquad\square$

Corollary 3.6 (Approximate geodesic convexity of F)**.** *Under the same assumption of the previous Theorem, let us choose $x, y \in D(F)$ with $\mathsf{d}_\ell(x, y) < \infty$, $\varepsilon > 0$, and points $x_0, \dots, x_N$, $N \in \mathbb{N}$, corresponding to a uniform partition $s_n = n/N$ of $[0, 1]$ such that*

$$
x_0 = x, \; x_N = y, \quad \mathsf{d}_\ell(x_i, x_{i+1}) \leq \frac{1}{N}\sqrt{\mathsf{d}_\ell^2(x, y) + \varepsilon^2} \quad 0 \leq i \leq (N - 1) . \tag{3.15}
$$

Then for all $t > 0$ one has

$$
F(\mathsf{S}_t x_n) \leq (1 - s_n)F(x) + s_n F(y) - \frac{K}{2}s_n(1 - s_n)\mathsf{d}_\ell^2(x, y) + \frac{\varepsilon^2}{2\mathrm{I}_K(t)}s_n(1 - s_n) . \tag{3.16}
$$

In particular, $(D(F), \mathsf{d}_\ell)$ and all sublevels of F are length spaces. If moreover the sublevels $\{F \leq c\}$, $c \in \mathbb{R}$, are complete in (X, d), then $\mathsf{d}_\ell = \bar{\mathsf{d}}_\ell$ on $D(F) \times D(F)$.

Proof. The inequality (3.16) can be obtained from the EVI_K property relative to d_ℓ (whose validity is ensured by Theorem 3.5) arguing as in [19, Thm. 3.2]. The equality $\mathsf{d}_\ell = \bar{\mathsf{d}}_\ell$ on the sublevels and then on $D(F)$ follows by the equality between the two length distances induced by complete distances. $\qquad\square$

An important application of the above result, that we will exploit in Section 11 (specifically, with $F = \mathrm{Ent}$ and τ equal to the weak $L^1(X, \mathfrak{m})$-topology) concerns the case when the completeness of the sublevels $\{F \leq c\}$ can be improved to compactness with respect to a Hausdorff topology τ on X for which d is lower semicontinuous. In this case it is not difficult to prove

that $(D(F), \mathsf{d}_\ell)$ is a *geodesic* space and that (see [18]) the EVI_K property relative to d_ℓ yields K-convexity of F, i.e. for every $x, y \in D(F)$ there exists a constant speed geodesic $x : [0, 1] \to X$ relative to d_ℓ connecting x to y such that

$$\mathsf{d}_\ell(x_s, x_t) = |t - s|\mathsf{d}_\ell(x, y), \quad F(x_t) \leq (1 - t)F(x) + tF(y) - \frac{K}{2}t(1 - t)\mathsf{d}_\ell^2(x, y) . \qquad (3.17)$$

4. EXTENDED METRIC-TOPOLOGICAL SPACES

We axiomatize metric-topological spaces by adding two (somehow competing) compatibility conditions between the distance d and the topology τ.

Definition 4.1 (Extended metric-topological spaces)**.** *Let (X, d) be an extended metric space and let τ be a Hausdorff topology in X. We say that (X, τ, d) is an extended metric-topological space if:*

    (a) *there exists a family of $(\tau \times \tau)$-continuous bounded semidistances $\mathsf{d}_i : X \times X \to [0, \infty)$, $i \in I$, with $\mathsf{d} = \sup_i \mathsf{d}_i$.*
    (b) *the topology τ is generated by the family of functions*

$$\mathrm{Lip}_b(X, \tau, \mathsf{d}) := \{f : X \to \mathbb{R} : \ f \text{ is bounded, } \mathsf{d}\text{-Lipschitz, } \tau\text{-continuous}\} . \qquad (4.1)$$

For every $L \geq 0$ we will also set

$$\mathrm{Lip}_{b,L}(X, \tau, \mathsf{d}) := \left\{f \in \mathrm{Lip}_b(X, \tau, \mathsf{d}) : \ |f(x) - f(y)| \leq L\mathsf{d}(x, y) \quad \forall \, x, y \in X\right\} \qquad (4.2)$$

Let us make a few comments on the above definition.

Boundedness of d_i. The boundedness assumption on d_i in (a) is clearly not restrictive, possibly replacing d_i by $\mathsf{d}_i \wedge n$ and enlarging the index set.

Directed families of distances. Possibly passing from the index set I to the collection of its finite subsets we can assume with no loss of generality that I is a directed set. We shall often make this assumption in the sequel.

Extended metric-topological structures generated by separating family of functions. If we don't take the extended metric structure as a starting point, a definition easily seen to be equivalent can be given starting from a class $\mathcal{A}$ of functions which separate the points of X; then τ is defined as the topology generated by $\mathcal{A}$ and the extended distance $\mathsf{d}(x, y)$ can be obtained by taking the supremum of $|f(x) - f(y)|$ as f runs in $\mathcal{A}$, as in (3.1). Notice that in this case we can take as topology τ the coarsest topology that makes all functions in $\mathcal{A}$ continuous, which is easily seen to be Hausdorff; since $\mathrm{Lip}_b(X, \tau, \mathsf{d})$ contains $\mathcal{A}$ by construction, it turns out that condition (a) above is satisfied and condition (b) is satisfied as well, with $\mathsf{d}_i(x, y) = |f_i(x) - f_i(y)|$, $\mathcal{A} = \{f_i\}_{i \in I}$. In addition, if the functions in $\mathcal{A}$ are already continuous for some preexisting topology τ_0 in X, this construction provides a topology τ coarser than τ_0.

Relations between τ and d. Notice that condition (a) yields

$$\mathsf{d} \quad \text{is } (\tau \times \tau)\text{-lower semicontinuous in } X \times X \qquad (4.3)$$

and for every net $(x_j)_{j \in J}$ and every $x \in X$

$$\mathsf{d}(x_j, x) \to 0 \text{ implies } x_j \to x \text{ w.r.t. the topology } \tau \qquad (4.4)$$

Indeed, one can use assumption (b) and observe that $f(x_j) \to f(x)$ for all $f \in \mathrm{Lip}_b(X, \tau, \mathsf{d})$.

Lipschitz functions generating τ. We might equivalently express condition (b) by assuming that there exists a smaller family $\mathcal{F} \subset \mathrm{Lip}(X, \tau, \mathsf{d})$ which generates the topology τ, i.e. for every net $(x_j)_{j \in J}$ in X and every $x \in X$

$$f(x_j) \to f(x) \text{ for all } f \in \mathcal{F} \text{ implies } x_j \to x \text{ w.r.t. the topology } \tau . \tag{4.5}$$

In fact, by suitably modifying the set of distances d_i without changing τ and d, we can obtain an equivalent characterization of extended metric-topological spaces:

Lemma 4.2. (X, τ, d) *is an extended metric-topological space according to Definition 4.1 if and only if* (X, d) *is an extended metric space and there exists a family of* $(\tau \times \tau)$-*continuous and bounded semidistances* d_i, $i \in I$, *such that* $\mathsf{d} = \sup_{i \in I} \mathsf{d}_i$ *and for every net* $(x_j)_{j \in J}$ *in* X *and* $x \in X$

$$\lim_{j \in J} x_j = x \quad w.r.t. \ \tau \qquad \Longleftrightarrow \qquad \lim_{j \in J} \mathsf{d}_i(x_j, x) = 0 \quad for \ every \ i \in I . \tag{4.6}$$

Proof. If (X, τ, d) is an extended metric-topological space according to Definition 4.1 we can simply consider the new collection of semidistances of the form

$$\mathsf{d}_f(x, y) := |f(x) - f(y)|, \qquad f \in \mathcal{F} = \mathrm{Lip}_{b,1}(X, \tau, \mathsf{d}) \tag{4.7}$$

Conversely, if a family of $(\tau \times \tau)$-continuous and bounded semidistances d_i, $i \in I$, satisfy $\mathsf{d} = \sup_{i \in I} \mathsf{d}_i$ and (4.6), then it is easy to check that $\mathcal{F} := \{\mathsf{d}_i(\cdot, z) : i \in I, \ z \in X\}$ is contained in $\mathrm{Lip}_b(X, \tau, \mathsf{d})$ and generates the topology τ. $\qquad\square$

Approximation of continuous functions. Every $f \in C_b(X)$ admits the useful representation formula

$$f(x) = \sup_{g \in L^-(f)} g(x) \quad \text{with} \quad L^-(f) := \Big\{ g \in \mathrm{Lip}_b(X, \tau, \mathsf{d}), \quad g \leq f \text{ in } X \Big\}$$
$$f(x) = \inf_{h \in L^+(f)} h(x) \quad \text{with} \quad L^+(f) := \Big\{ h \in \mathrm{Lip}_b(X, \tau, \mathsf{d}), \quad h \geq f \text{ in } X \Big\} \tag{4.8}$$

which can be proved by passing to the limit with respect to $(i, n) \in I \times \mathbb{N}$ in the inf and sup regularizations of f

$$g_{i,n}(x) := \inf_{y \in X} f(y) + n\mathsf{d}_i(x, y), \qquad h_{i,n}(x) := \sup_{y \in X} f(y) - n\mathsf{d}_i(x, y) \tag{4.9}$$

associated to any family of $(\tau \times \tau)$-continuous distances d_i satisfying (4.6).

Properties (4.3) and (4.4) show that Definition 4.1 is consistent with the axiomatization of extended metric spaces proposed in [3]. Notice however that we assume neither that τ is Polish (by working directly with Radon measures) nor that (X, d) is complete.

Completion. Thanks to (4.1), extended metric-topological spaces behave well w.r.t. completion: denoting by $(\widetilde{X}, \widetilde{\mathsf{d}})$ the abstract completion of (X, d), every function f in $\mathrm{Lip}_b(X, \tau, \mathsf{d})$ admits a unique Lipschitz extension $\widetilde{f}$ to $\widetilde{X}$ and we can thus introduce the topology $\widetilde{\tau}$ generated by $\{\widetilde{f} : f \in \mathrm{Lip}_b(X, \tau, \mathsf{d})\}$. It is not difficult to check that $(\widetilde{X}, \widetilde{\tau}, \widetilde{\mathsf{d}})$ is an extended metric-topological space and that τ is the restriction of $\widetilde{\tau}$ to X.

A canonical compactification. Consider a set F and the space $\mathcal{X} := \mathbb{R}^F$ endowed with the topology $\tau_{\mathcal{X}}$ of pointwise convergence. On $\mathcal{X}$ we consider the extended distance

$$\mathsf{d}_{\mathcal{X}}(\boldsymbol{x}, \boldsymbol{y}) := \sup_{f \in F} |x_f - y_f| \quad \text{with} \quad \boldsymbol{x} = (x_f)_{f \in F}, \ \boldsymbol{y} = (y_f)_{f \in F} \in \mathcal{X} .$$

$(\mathcal{X}, \tau_{\mathcal{X}}, \mathsf{d}_{\mathcal{X}})$ provides a natural class of example of extended metric-topological space depending on the index set F. Extended metric-topological spaces can always be embedded in a compact subset of some space $\mathcal{X} = \mathbb{R}^F$, preserving the metric-topological structure. As index set F and embedding ι we can always choose

$$F = \mathrm{Lip}_{b,1}(X, \tau, \mathsf{d}), \quad \iota : X \to \mathbb{R}^F, \quad \iota(x)_f := f(x) \quad f \in F \tag{4.10}$$

as is typical for the Stone-Čech compactification of completely regular spaces. The proof of the following result is standard.

Lemma 4.3. *If (X, τ, d) is an extended metric-topological space according to Definition 4.1, then the map ι in (4.10) is an homeomorphism of X with $\widetilde{X} = \iota(X) \subset \mathcal{X} = \mathbb{R}^F$, it is an isometry from (X, d) to $(\widetilde{X}, \mathsf{d}_{\mathcal{X}})$, and $\widetilde{X}$ has a $\tau_{\mathcal{X}}$-compact closure in $\mathcal{X}$.*

One more important consequence, from the measure-theoretic point of view, is the following result concerning the complete regularity, according to (2.5), of metric-topological spaces.

Lemma 4.4 (Complete regularity)**.** *Any extended metric-topological space (X, τ, d) is completely regular.*

Proof. Let $x_0 \in U \in \tau$ with $U \cap F = \emptyset$. By condition (b), we can assume that U has the form $\cap_j \{f_j > 0\}$ for some finite family of functions $f_j \in \mathrm{Lip}_b(X, \tau, \mathsf{d})$. Then the function $g = \min_j f_j^+$ is null on F and $g(x_0) > 0$. $\qquad\square$

A first consequence of the $(\tau \times \tau)$-lower semicontinuity of d, technically relevant for us, is the τ-lower semicontinuity of the function

$$\mathsf{d}_K(x) := \min_{y \in K} \mathsf{d}(x, y) \tag{4.11}$$

for any compact set $K \subset X$. The function d_K, whose 0 level set is K, provides in our context an analogous of the *perfect* regularity condition in Topology. Notice also that d_K is the monotone limit of the τ-continuous functions

$$\mathsf{d}_K^i(x) := \min_{y \in K} \mathsf{d}_i(x, y) . \tag{4.12}$$

Let us now introduce some additional concepts where the topological and the metric structure interact. We denote by $\mathscr{B}(\tau, \mathsf{d})$ the σ-algebra generated by $\mathrm{Lip}_b(X, \tau, \mathsf{d})$; obviously $\mathscr{B}(\tau, \mathsf{d}) \subset \mathscr{B}(\tau)$, but a kind of converse is provided by the following lemma.

Lemma 4.5. *In an extended metric-topological space (X, τ, d), any $\mu \in \mathscr{P}(X)$ is uniquely determined by its values on $\mathscr{B}(\tau, \mathsf{d})$. In addition $\cup_i \mathrm{Lip}_b(X, \tau, \mathsf{d}_i)$ separates points of X and it is dense in $L^2(X, \mathfrak{m})$. In particular $\mathrm{Lip}_b(X, \tau, \mathsf{d})$ separates points of X and it is dense in $L^2(X, \mathfrak{m})$.*

Proof. For the first statement it is sufficient to remark that the sets $\{\mathsf{d}_K^i > 0\}$ are τ-open, belong to $\mathscr{B}(\tau, \mathsf{d})$ and monotonically converge to $\{\mathsf{d}_K > 0\} = X \setminus K$. Then the Radon property of μ ensures that $\mu(X \setminus K) = \lim_i \mu(\{\mathsf{d}_K^i > 0\})$. Since $\{\mathsf{d}(\cdot, x)\}_{x \in X}$ separates points of X, the family $\{\mathsf{d}_i(\cdot, x)\}_{x \in X, i \in I}$ is contained in $\mathrm{Lip}_b(X, \tau, \mathsf{d})$ and separates points as well. To prove the density, it is sufficient to show the implication

$$\int f\phi \, \mathrm{d}\mathfrak{m} = 0 \qquad \forall \phi \in \bigcup_{i \in I} \mathrm{Lip}_b(X, \tau, \mathsf{d}_i) \qquad \Longrightarrow \qquad f = 0 .$$

Clearly $\int f \, \mathrm{d}\mathfrak{m} = 0$ and, arguing by contradiction, it is not restrictive to assume $\int |f| \, \mathrm{d}\mathfrak{m} = 2$. Splitting f in positive and negative part we can consider $\mu^\pm := f^\pm \mathfrak{m} \in \mathscr{P}(X)$ to obtain that $\mu^+ = \mu^-$, which is a contradiction. $\qquad\square$

On X^D we put the product topology $\tau^{\otimes D}$, i.e. the coarsest topology making all e_t continuous. We denote for simplicity this topology by τ^*. Notice that (X^D, τ^*) is Hausdorff, because $\{\mathrm{e}_t\}_{t \in D}$ separates points of X^D; if (X, τ) is completely regular then (X^D, τ^*) is completely regular as well, since it is a product of completely regular spaces [35, Thm. 33.2].

By continuity, the push-forward operator induced by e_t maps Radon measures in X^D to Radon measures in X. Let us now check that the class $UC(D; (X, \mathsf{d}))$ of d-uniformly continuous paths $\eta : D \to X$ is a Borel subset of X^D. Indeed, it is easily seen that its complement is described by

$$\bigcup_{k=1}^{\infty} \bigcap_{\ell=1}^{\infty} \bigcup_{i \in I} \left\{ \eta : \ \exists s, t \in D \text{ with } |s - t| \leq \frac{1}{\ell}, \ \mathsf{d}_i(\eta(s), \eta(t)) > \frac{1}{k} \right\} .$$

On the other hand, for k, ℓ and i fixed, the complement of the set above is

$$\bigcap_{s, t \in D, \ell |s - t| \leq 1} \left\{ \eta : \ \mathsf{d}_i(\eta(s), \eta(t)) \leq \frac{1}{k} \right\}$$

and therefore is closed in X^D (by the continuity of $(\mathsf{e}_s, \mathsf{e}_t) : X^D \to X \times X$).

It is easy to check that $\mathsf{d} \circ (\mathsf{e}_s, \mathsf{e}_t) = \sup_i \mathsf{d}_i \circ (\mathsf{e}_s, \mathsf{e}_t)$ is τ^*-lower semicontinuous in X^D for all $s, t \in D$. This can be used to prove that $\mathscr{A}_p$ is τ^*-lower semicontinuous in X^D and, in particular, that $AC^p(D; (X, \mathsf{d}))$ is a Borel subset of X^D, more precisely a countable union of closed sets.

Theorem 4.6 (Compactness of probabilities in X^D). *Assume that $p \in (1, \infty)$, $T > 0$ and $D \subset \mathbb{R}$ is countable. Let $(\boldsymbol{\eta}_i)_{i \in I}$ be a net of Radon probability measures in (X^D, τ^*) such that for all $t \in D$ the family $\{(\mathsf{e}_t)_\# \boldsymbol{\eta}_i\}_{i \in I}$ is equi-tight in $\mathscr{P}(X)$.*
Then $\{\boldsymbol{\eta}_i\}_{i \in I}$ has limit points $\boldsymbol{\eta} \in \mathscr{P}(X^D)$ in the weak topology induced by $C_b(X^D)$ and any such limit point, along a subnet $\beta : L \to I$, satisfies:

$$\int \mathscr{A}_p(\eta, D') \, \mathrm{d}\boldsymbol{\eta}(\eta) \leq \liminf_{\ell \in L} \int \mathscr{A}_p(\eta, D') \, \mathrm{d}\boldsymbol{\eta}_{\beta(\ell)}(\eta) \qquad \forall D' \subset D . \tag{4.13}$$

Proof. The family $\{\boldsymbol{\eta}_i\}_{i \in I}$ is equi-tight: indeed, enumerating by $\{d_k\}_{k \in \mathbb{N}}$ the elements of D, it suffices to find compact sets $K_{k,n}$ such that $\sup_i (\mathsf{e}_{d_k})_\# \boldsymbol{\eta}_i(X \setminus K_{k,n}) \leq 2^{-k-n}$ and to consider the sets

$$\Gamma_n := \bigcap_{k=0}^{\infty} \big\{ \eta \in X^D : \ \eta(d_k) \in K_{k,n} \big\}$$

which are compact in X^D and satisfy $\sup_i \boldsymbol{\eta}_i(X^D \setminus \Gamma_n) \leq 2^{1-n}$. It follows that we can apply Theorem 2.1.

In order to prove (4.13) for a limit point $\boldsymbol{\eta} \in \mathscr{P}(X)$ we use the continuity of $\eta \mapsto \mathsf{d}_i(\eta(s), \eta(t))$ and the Radon property of $\boldsymbol{\eta}$ to get

$$\int \mathsf{d}^p(\eta(s), \eta(t)) \, \mathrm{d}\boldsymbol{\eta}(\eta) \leq \liminf_{\ell \in L} \int \mathsf{d}^p(\eta(s), \eta(t)) \, \mathrm{d}\boldsymbol{\eta}_{\beta(\ell)}(\eta) .$$

Then, given $D' \subset D$, from the superadditivity of $\liminf$ we obtain, for all choices of $t_0, \ldots, t_n \in D'$,

$$\int \sum_{j=0}^{n-1} \frac{\mathsf{d}^p(\eta(t_{j+1}), \eta(t_j))}{(t_{j+1} - t_j)^{p-1}} \, \mathrm{d}\boldsymbol{\eta}(\eta) \ \leq \ \liminf_{\ell \in L} \int \sum_{j=0}^{n-1} \frac{\mathsf{d}^p(\eta(t_{j+1}), \eta(t_j))}{(t_{j+1} - t_j)^{p-1}} \, \mathrm{d}\boldsymbol{\eta}_{\beta(\ell)}(\eta)$$

$$\leq \ \liminf_{\ell \in L} \int \mathscr{A}_p(\eta, D') \, \mathrm{d}\boldsymbol{\eta}_{\beta(\ell)}(\eta) .$$

Since the sums $\sum_i \mathsf{d}^p(\eta(t_{j+1}), \eta(t_j)) / (t_{j+1} - t_j)^{p-1}$ become larger as the partition gets finer, we can use once more the fact that $\boldsymbol{\eta}$ is Radon to conclude. $\qquad\square$

Finally, we can add a measure structure in extended metric-topological spaces as follows.

Definition 4.7 (Extended metric measure space). *We say that $(X, \tau, \mathsf{d}, \mathfrak{m})$ is an extended metric measure space if:*
  (a) *(X, τ, d) is an extended metric-topological space;*
  (b) *$\mathfrak{m} \in \mathscr{P}(X)$, i.e. $\mathfrak{m}$ is a Radon probability measure in $(X, \mathscr{B}(\tau))$.*

5. The Wasserstein space over an extended metric-topological space

Throughout this section (X, τ, d) is an extended metric-topological space.

5.1. **The extended Wasserstein distance between Radon measures.** In the class $\mathscr{P}(X)$ of Radon probability measures in X, we define the (quadratic) Wasserstein extended distance $W_\mathsf{d}(\mu, \nu)$ by

$$W_\mathsf{d}^2(\mu, \nu) := \inf \left\{ \int_{X \times X} \mathsf{d}^2 \, \mathrm{d}\boldsymbol{\pi} : \ \boldsymbol{\pi} \in \Gamma(\mu, \nu) \right\} \tag{5.1}$$

where we recall that $\Gamma(\mu, \nu)$ is the class of admissible transport plans between μ and ν, see (2.8).

In our context, since (X, τ) is completely regular thanks to Lemma 4.4, we know that narrow and weak topology coincide. From now on, unless otherwise stated, by weak convergence we always mean convergence in the duality with C_b, the corresponding topology will be denoted by $\tau_\mathscr{P}$.

The lower semicontinuity of the cost d^2 and the tightness of the marginals ensure respectively lower semicontinuity of the transportation cost and compactness w.r.t. weak convergence of the class of the admissible transport plans $\Gamma(\mu, \nu)$, hence existence of optimal plans. We provide a more general statement in the next theorem.

Theorem 5.1 (Compactness and joint lower semicontinuity). *Let I be a directed set and assume that $\mu_i, \nu_i \in \mathscr{P}(X)$ weakly converge to $\mu, \nu \in \mathscr{P}(X)$ respectively. Then, for any choice of $\boldsymbol{\pi}_i \in \Gamma(\mu_i, \nu_i)$, one has:*

(a) *the net $(\boldsymbol{\pi}_i)_{i \in I}$ has limit points w.r.t. weak convergence and any limit point $\boldsymbol{\pi}$ belongs to $\Gamma(\mu, \nu)$;*

(b) *if $\boldsymbol{\pi}_i$ weakly converges to $\boldsymbol{\pi}$, and $\mathsf{c}_i : X \times X \to [0, \infty]$ is a monotone family of $(\tau \times \tau)$-lower semicontinuous functions, then*

$$\liminf_{i \in I} \int \mathsf{c}_i \, \mathrm{d}\boldsymbol{\pi}_i \geq \int \mathsf{c} \, \mathrm{d}\boldsymbol{\pi} \qquad \text{with } \mathsf{c} := \sup_i \mathsf{c}_i \, ;$$

(c) *if $\mu_i = \mu$, $\nu_i = \nu$, and if $\mathsf{d}_i : X \times X \to [0, \infty]$ is a monotone family of $(\tau \times \tau)$-continuous distances with $\lim_i \mathsf{d}_i = \mathsf{d}$, then W_{d_i} monotonically converges to W_{d}.*

Proof. Statement (a) is a direct consequence of Theorem 2.1, since $(X \times X, \tau \times \tau)$ is completely regular. Statement (b) follows by (2.3). Statement (c) follows by (b) and (a) with $\mu_i = \mu$, $\nu_i = \nu$, choosing optimal plans $\boldsymbol{\pi}_i$ relative to d_i^2 and extracting a weakly convergent subnet from the $\boldsymbol{\pi}_i$. $\qquad \square$

We claim that the duality formula

$$\frac{1}{2} W_{\mathsf{d}}^2(\mu, \nu) = \sup_{\phi \in \mathcal{F}} \int Q_1 \phi \, \mathrm{d}\nu - \int \phi \, \mathrm{d}\mu \tag{5.2}$$

holds, where $\mathcal{F}$ is defined in (2.11) and $Q_t \phi$ is defined by the Hopf-Lax formula

$$Q_t \phi(y) := \inf_{x \in X} \phi(x) + \frac{1}{2t} \mathsf{d}^2(x, y) \, . \tag{5.3}$$

In fact, from the very definition of $\mathcal{F}$ in (2.11) it follows that

$$Q_t \phi(y) = C \wedge \min_{x \in K} \left(\phi(x) + \frac{1}{2t} \mathsf{d}^2(x, y) \right) \tag{5.4}$$

for some compact set $K \subset X$ with $\phi \in C(K)$ and $C \geq \max_K \phi$, hence $Q_t \phi$ are τ-lower semicontinuous and we can replace $\int_* Q_1 \phi \, \mathrm{d}\nu$ with $\int Q_1 \phi \, \mathrm{d}\nu$. In addition, the compactness of K ensures that

$$\min_{x \in K} \left(\phi(x) + \frac{1}{2t} \mathsf{d}_i^2(x, y) \right) \uparrow \min_{x \in K} \left(\phi(x) + \frac{1}{2t} \mathsf{d}^2(x, y) \right)$$

hence from Proposition 2.3 (which deals with bounded cost functions) with $\frac{1}{2} \mathsf{d}_i^2$ and statement (c) of the previous theorem we obtain (5.2).

We will occasionally use also the extended Wasserstein distance $W_{\mathsf{d},1}$ in $\mathscr{P}(X)$ obtained by minimizing $\int \mathsf{d} \, \mathrm{d}\boldsymbol{\pi}$ in the class of admissible transport plans, and the corresponding duality formula

$$W_{\mathsf{d},1}(\mu, \nu) = \sup \left\{ \int f \, \mathrm{d}(\mu - \nu) : \ f \in \mathrm{Lip}_b(X, \tau, \mathsf{d}), \ \mathrm{Lip}(f) \leq 1 \right\} \, . \tag{5.5}$$

Also the proof of (5.5) can be obtained from Proposition 2.3 with $\mathsf{c} = \mathsf{d}_i$, considering the pairs $(-\phi_i^{\mathsf{c}}, \phi_i^{\mathsf{c}})$ with

$$\phi_i^{\mathsf{c}}(x) := C \wedge \min_{x \in K} \left(\phi(x) + \mathsf{d}_i(x, y) \right) \in \mathrm{Lip}_b(X, \tau, \mathsf{d})$$

and $\phi \in \mathcal{F}$, $K \subset X$ compact set as in (2.11), $C \geq \max_K \phi$.

Proposition 5.2 (W_{d} convergence implies $\tau_{\mathscr{P}}$ convergence). *A net $\{\mu_j\}_{j \in J}$ weakly converges to μ in $\mathscr{P}(X)$ if and only if*

$$\lim_{j \in J} \int f \, \mathrm{d}\mu_j = \int f \, \mathrm{d}\mu \quad \text{for every } f \in \mathrm{Lip}(X, \tau, \mathsf{d}) \, . \tag{5.6}$$

In particular, if $\{\mu_j\}_{j \in J}$ converges to μ w.r.t. W_d, then it also converges w.r.t. the weak topology induced by $C_b(X)$.

Proof. Let us first prove that (5.6) is sufficient to prove the weak convergence of μ_j in $\mathscr{P}(X)$ (the converse implication is trivial). If $f \in C_b(X)$ we consider the directed set $L^-(f)$ defined in (4.8), obtaining

$$\liminf_{j \in J} \int f \, \mathrm{d}\mu_j \geq \lim_{j \in J} \int g \, \mathrm{d}\mu_j = \int g \, \mathrm{d}\mu \quad \text{for every } g \in L^-(f)$$

so that

$$\liminf_{j \in J} \int f \, \mathrm{d}\mu_j \geq \sup_{g \in L^-(f)} \int g \, \mathrm{d}\mu = \lim_{g \in L^-(f)} \int g \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu$$

since $L^-(f)$ is a directed set with respect to the natural ordering of functions and μ is a Radon measure. Changing f in $-f$ we get the opposite inequality for the $\limsup$, thus proving that $\lim_{j \in J} \int f \, \mathrm{d}\mu_j = \int f \, \mathrm{d}\mu$.

If $\{\mu_j\}_{j \in J}$ is a net convergent w.r.t. W_d to $\mu \in \mathscr{P}(X)$ the inequality (ensured by $W_{\mathsf{d},1} \leq W_\mathsf{d}$ and (5.5))

$$\left| \int f \, \mathrm{d}\mu - \int f \, \mathrm{d}\mu_j \right| \leq \mathrm{Lip}(f) W_\mathsf{d}(\mu_j, \mu) \tag{5.7}$$

yields (5.6) and therefore the weak convergence of μ_j. $\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

The following result shows the flexibility of our axiomatization: the extended metric-topological structure can be lifted from the space X to the space of probabilities on X.

Proposition 5.3 (Extended metric-topological structure on $\mathscr{P}(X)$). *$(\mathscr{P}(X), \tau_{\mathscr{P}}, W_\mathsf{d})$ is an extended metric-topological space.*

Proof. Let us first show that condition (a) of Definition 4.1 is satisfied.

We consider the set $F = \mathrm{Lip}_{b,1}(X, \tau, \mathsf{d})$ and we denote by I the collection (directed set) of the finite subsets of F ordered by inclusion. For every $i \in I$ we can set

$$M_i := \sup_{x \in X} \sup_{f \in i} |f(x)| \qquad \mathsf{d}_i(x, y) := \sup_{f \in i} |f(x) - f(y)| \, .$$

Since we already noticed that $W_{\mathsf{d}_i} \uparrow W_\mathsf{d}$, it remains to prove that each semidistance W_{d_i} is $(\tau_{\mathscr{P}} \times \tau_{\mathscr{P}})$-continuous. So, for a fixed $i = \{f_1, \ldots, f_N\} \subset F$ we consider the τ-continuous map $\kappa : X \to X_i = [-M_i, M_i]^N$ and the distance δ in X_i defined by

$$\kappa(x) := (f_1(x), \ldots, f_N(x)), \quad \delta(\boldsymbol{x}, \boldsymbol{y}) := \sup_{n=1,\ldots,N} |x_n - y_n|, \quad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^N \, ,$$

so that κ is 1-Lipschitz. Since in $\mathscr{P}(X_i)$ the weak topology coincides with the topology induced by the Wasserstein distance W_δ, $\kappa_\sharp$ is continuous from $(\mathscr{P}(X), \tau_{\mathscr{P}})$ to $(\mathscr{P}(X_i), W_\delta)$; it is therefore sufficient to prove that

$$W_{\mathsf{d}_i}(\mu, \nu) \leq W_\delta(\kappa_\sharp \mu, \kappa_\sharp \nu) \quad \text{for every } \mu, \nu \in \mathscr{P}(X) \tag{5.8}$$

which in fact yields the equality, since the opposite inequality is trivial. To prove (5.8) we apply Proposition 2.3 with $\mathsf{c} = \mathsf{d}_i^2$ to find a sequence of uniformly bounded Borel functions ϕ_n, ψ_n such that

$$\psi_n(y) - \phi_n(x) \leq \mathsf{d}_i^2(x, y) \quad \text{for every } x, y \in X, \quad W_{\mathsf{d}_i}^2(\mu, \nu) = \lim_{n \to \infty} \int \psi_n \, \mathrm{d}\nu - \int \phi_n \, \mathrm{d}\mu \, .$$

Possibly replacing ψ_n by ϕ_n^{c} and ϕ_n by ϕ_n^{cc} and using the fact that c-concave functions are d_i-Lipschitz (and therefore τ-continuous) we may assume that $\phi_n, \psi_n \in \mathrm{Lip}_b(X, \mathsf{d}_i)$. We can then define functions $\widetilde{\phi}_n, \widetilde{\psi}_n : \kappa(X_i) \to \mathbb{R}$ by

$$\widetilde{\phi}_n(\kappa(x)) := \phi_n(x), \quad \widetilde{\psi}_n(\kappa(x)) := \psi_n(x) \quad x \in X$$

the definition being consistent since $\kappa(x) = \kappa(y)$ yields $\mathsf{d}_i(x, y) = 0$ and therefore $\phi_n(x) = \phi_n(y)$ and $\psi_n(x) = \psi_n(y)$. Since moreover $\delta(\kappa(x), \kappa(y)) = \mathsf{d}_i(x, y)$, we can easily prove that $\widetilde{\phi}_n, \widetilde{\psi}_n$ are Lipschitz continuous w.r.t. δ in $\kappa(X)$, and therefore they admit a unique Lipschitz continuous

extension (still denoted by $\widetilde{\phi}_n$, $\widetilde{\psi}_n$) to the compact set $\overline{\kappa(X)}$. Notice that both the supports $\mathrm{supp}(\kappa_\sharp \mu)$ and $\mathrm{supp}(\kappa_\sharp \nu)$ in X_i are contained in $\overline{\kappa(X)}$. Moreover the relation

$$\widetilde{\psi}_n(y) - \widetilde{\phi}_n(x) \le \delta(x,y) \quad \text{for every } x, y \in \kappa(X)$$

extends by continuity to $\overline{\kappa(X)}$, so that

$$W_\delta^2(\kappa_\sharp \mu, \kappa_\sharp \nu) \ge \int \widetilde{\psi}_n \, \mathrm{d}\kappa_\sharp \nu - \int \widetilde{\phi}_n \, \mathrm{d}\kappa_\sharp \mu = \int \psi_n \, \mathrm{d}\nu - \int \phi_n \, \mathrm{d}\mu$$

proving (5.8).

In order to prove that also condition (b) of Definition 4.1 is satisfied, we just observe that the family $\mathcal{F}$ of real functions on $\mathscr{P}(X)$ of the form

$$f[\mu] : \mu \mapsto \int f \, \mathrm{d}\mu, \quad f \in \mathrm{Lip}_{b,1}(X, \tau, \mathsf{d})$$

is included in $\mathrm{Lip}_{b,1}(\mathscr{P}(X), \tau_{\mathscr{P}}, W_{\mathsf{d}})$ thanks to the very definition of weak convergence and to (5.7). On the other hand, Proposition 5.2 shows that $\tau_{\mathscr{P}}$ is induced by $\mathcal{F}$. $\qquad \square$

Proposition 5.4 (Completeness of $(\mathscr{P}(X), W_{\mathsf{d}})$). *If (X, τ, d) is an extended metric-topological space and (X, d) is complete, then also $(\mathscr{P}(X), W_{\mathsf{d}})$ is complete.*

Proof. Let $(\mu_n)_{n\in\mathbb{N}} \subset \mathscr{P}(X)$ be a sequence satisfying $\sum_{n=1}^\infty W_{\mathsf{d}}(\mu_n, \mu_{n+1}) < \infty$. We argue as in the proof of [2, Prop. 7.1.5]: if $\boldsymbol{\pi}_n \in \Gamma(\mu_n, \mu_{n+1})$ are optimal transport plans, by applying Lemma 2.2 we can find a Radon measure $\boldsymbol{\pi} \in \mathscr{P}(\boldsymbol{X})$, $\boldsymbol{X} = X^{\mathbb{N}}$, such that $(\mathsf{p}^n, \mathsf{p}^{n+1})_\sharp \boldsymbol{\pi} = \boldsymbol{\pi}_n$, $n \in \mathbb{N}$.

We thus have $\sum_{n=1}^N \int \mathsf{d}(\mathsf{p}^n, \mathsf{p}^{n+1}) \, \mathrm{d}\boldsymbol{\pi} < \infty$ so that the sequence $n \mapsto \mathsf{p}^n(\boldsymbol{x})$ is a Cauchy sequence for $\boldsymbol{\pi}$-a.e. $\boldsymbol{x} \in \boldsymbol{X}$. Denoting by $\mathsf{p}(\boldsymbol{x})$ its pointwise limit, and applying Egoroff Theorem [39, Thm. 6, p. 28] we can find for every $\varepsilon > 0$ compact sets $K_\varepsilon \subset \boldsymbol{X}$ with $\boldsymbol{\pi}(\boldsymbol{X} \setminus K_\varepsilon) < \varepsilon$ such that the restrictions of p^n to K_ε converge uniformly, i.e.

$$\lim_{n\to\infty} \sup_{\boldsymbol{x} \in K_\varepsilon} \mathsf{d}(\mathsf{p}^n(\boldsymbol{x}), \mathsf{p}(\boldsymbol{x})) = 0 \,. \tag{5.9}$$

Let us prove that p is a Lusin $\boldsymbol{\pi}$-measurable map by showing that the restriction of p to each K_ε is continuous: the latter property will be a consequence of the fact that d-uniform limit of τ-continuous maps is τ-continuous. In fact, for every $\boldsymbol{x} \in K_\varepsilon$ and every τ-neighbourhood V of $\mathsf{p}(\boldsymbol{x})$ in X, we may find functions $f_j \in \mathrm{Lip}_{b,1}(X, \tau, \mathsf{d})$, $j = 1, \ldots, J$, and $\delta > 0$ such that

$$|f_j(\boldsymbol{x}) - f_j(\boldsymbol{y})| \le \delta \quad \text{for every } j = 1, \ldots, J \qquad \Rightarrow \qquad \boldsymbol{y} \in V \,.$$

If $n \in \mathbb{N}$ is sufficiently big so that $\sup_{\boldsymbol{z} \in K_\varepsilon} \mathsf{d}(\mathsf{p}^n(\boldsymbol{z}), \mathsf{p}(\boldsymbol{z})) < \delta/3$ and U is any neighbourhood in $\boldsymbol{X}$ such that every $\boldsymbol{y} \in U$ satisfies $|f_j(\mathsf{p}^n(\boldsymbol{x})) - f_j(\mathsf{p}^n(\boldsymbol{y}))| < \delta/3$ for every $j = 1, \ldots, N$, we conclude that for every $\boldsymbol{y} \in U \cap K_\varepsilon$

$$|f_j(\mathsf{p}(\boldsymbol{x})) - f_j(\mathsf{p}(\boldsymbol{y}))| \le 2 \sup_{\boldsymbol{z} \in K_\varepsilon} \mathsf{d}(\mathsf{p}(\boldsymbol{z})), \mathsf{p}^n(\boldsymbol{z})) + |f_j(\mathsf{p}^n(\boldsymbol{x})) - f_j(\mathsf{p}^n(\boldsymbol{y}))|$$

$$\le \delta \qquad \text{for every } j = 1, \ldots, N \,,$$

so that $\mathsf{p}(\boldsymbol{y}) \in V$.

Being p a Lusin $\boldsymbol{\pi}$-measurable map, $\mu := \mathsf{p}_\sharp \boldsymbol{\pi}$ is a Radon measure in X and

$$W_{\mathsf{d}}(\mu_n, \mu) \le \Big(\int \mathsf{d}^2(\mathsf{p}^n, \mathsf{p}) \, \mathrm{d}\boldsymbol{\pi} \Big)^{1/2} \le \sum_{m=n}^\infty \Big(\int \mathsf{d}^2(\mathsf{p}^m, \mathsf{p}^{m+1}) \, \mathrm{d}\boldsymbol{\pi} \Big)^{1/2} = \sum_{m=n}^\infty W_{\mathsf{d}}(\mu_m, \mu_{m+1}) \,,$$

which shows that $\lim_{n\to\infty} W_{\mathsf{d}}(\mu_n, \mu) = 0$. $\qquad \square$

5.2. **The superposition principle for extended metric-topological spaces.** The next proposition is a small variant of the superposition principle recently proved in [34, Thm. 3.1] for extended metric measure spaces. We provide a slightly different proof, since no Polish assumption on (X, τ) is made here, only the complete regularity of τ following by Lemma 4.4 plays a role.

Proposition 5.5 (Superposition). *Assume that (X, d) is complete and let $\mu_t \in AC^2([0, T]; (\mathscr{P}(X), W_{\mathsf{d}}))$. Then there exists $\boldsymbol{\eta} \in \mathscr{P}(X^{[0,T]})$ concentrated on $AC^2([0, T]; (X, \mathsf{d}))$ with $(\mathsf{e}_t)_{\#}\boldsymbol{\eta} = \mu_t$ for all $t \in [0, T]$ and*

$$\int |\dot{\eta}(t)|^2 \, \mathrm{d}\boldsymbol{\eta}(\eta) = |\dot{\mu}_t|^2 \qquad \text{for a.e. } t \in (0, T) . \tag{5.10}$$

Proof. Let us assume for simplicity $T = 1$. We set $D = \cup_n D_n$ with $D_n = \{j/2^n : 0 \le j \le 2^n\}$.
Step 1. We build $\boldsymbol{\sigma} \in \mathscr{P}(X^D)$ concentrated on $AC^2(D; (X, \mathsf{d}))$ satisfying $(\mathsf{e}_t)_{\#}\boldsymbol{\sigma} = \mu_t$ for all $t \in D$ and

$$\int \sum_{k=i}^{j-1} \frac{|\eta(k/2^n) - \eta((k+1)/2^n)|^2}{2^{-n}} \, \mathrm{d}\boldsymbol{\sigma}(\eta) \le \int_{i/2^n}^{j/2^n} |\dot{\mu}_r|^p \, \mathrm{d}r \quad \forall i, j \in \{0, \ldots, 2^n\}, \ i < j, \ n \ge 1 . \tag{5.11}$$

To this aim, taking Theorem 4.6, (4.13) and the fact that $D_n \uparrow D$ into account, it is sufficient to build a family of approximations $\boldsymbol{\sigma}_n \in \mathscr{P}(X^D)$ satisfying $(\mathsf{e}_t)_{\#}\boldsymbol{\sigma}_n = \mu_t$ for all $t \in D_n$ and

$$\int \sum_{k=i}^{j-1} \frac{|\eta(k/2^n) - \eta((k+1)/2^n)|^2}{2^{-n}} \, \mathrm{d}\boldsymbol{\sigma}_n(\eta) \le \int_{i/2^n}^{j/2^n} |\dot{\mu}_r|^2 \, \mathrm{d}r \qquad \forall i, j \in \{0, \ldots, 2^n\}, \ i < j . \tag{5.12}$$

The construction of $\boldsymbol{\sigma}_n$ is a simple application of Lemma 2.2: it suffices to choose optimal plans $\boldsymbol{\pi}_i$ from $\mu_{i/2^n}$ to $\mu_{(i+1)/2^n}$, $0 \le i \le 2^n$, and then use the lemma to find a Radon probability measure $\boldsymbol{\pi}$ in X^{2^n+1} having $\boldsymbol{\pi}_i$ as double marginals. Then one can define $\boldsymbol{\sigma}_n \in \mathscr{P}(X^D)$ as the push forward of $\boldsymbol{\pi}$ via the continuous map from X^{2^n+1} to X^D defined by

$$(x_0, \ldots, x_n) \mapsto \eta(t) = \begin{cases} x_i & \text{if } \frac{i}{2^n} \le t < \frac{i+1}{2^n}, \ 0 \le i \le n - 1 \\ x_n & \text{if } t = 1 . \end{cases}$$

Step 2. Given $\boldsymbol{\sigma}$ as in Step 1, we notice that $\boldsymbol{\sigma}$ is concentrated on the union of the closed sets $\Gamma_k = \{\eta : \mathscr{A}_2(\eta, D) \le k\}$. Since (X, d) is complete we can consider the extension map $\mathsf{ext} : \cup_k \Gamma_k \to AC^2([0, 1]; (X, \mathsf{d}))$ and build $\boldsymbol{\eta}$ as the image under $\boldsymbol{\sigma}$ of ext. To show that $\boldsymbol{\eta}$ is well defined and it is a Radon measure in $\mathscr{P}(X^{[0,T]})$ we need to show that ext (arbitrarily defined out of $\cup_k \Gamma_k$) is Lusin $\boldsymbol{\sigma}$-measurable. To this purpose, it is sufficient to prove that $\mathsf{ext} : \Gamma_k \to X^{[0,T]}$ is continuous. Let $(\eta_i)_{i \in I}$ be a net in Γ_k convergent to η and let $\widetilde{\eta}_i, \widetilde{\eta}$ be the corresponding extensions to $[0, 1]$. By the definition of product topology we need only to prove that $\widetilde{\eta}_i(t) \to \widetilde{\eta}(t)$ in (X, τ) for all $t \in [0, 1]$. Since τ is generated by $\mathrm{Lip}_b(X, \tau, \mathsf{d})$ we need only to prove that $f(\widetilde{\eta}_i(t)) \to f(\widetilde{\eta}(t))$ for all $f \in \mathrm{Lip}_b(X, \tau, \mathsf{d})$. This is trivial if $t \in D$ (because $\eta_i \to \eta$ in X^D), in the general case one can use the estimate

$$|f(\widetilde{\eta}_i(t)) - f(\widetilde{\eta}_i(s))| \le \mathrm{Lip}(f)\mathsf{d}(\widetilde{\eta}_i(t), \widetilde{\eta}_i(s)) \le \mathrm{Lip}(f)\sqrt{k|t - s|}$$

and the analogous one for $\widetilde{\eta}$ to conclude.

Having proved that $\boldsymbol{\eta}$ is well defined $\mathscr{P}(X^{[0,T]})$ and it is concentrated on the Borel set $AC^2([0, 1]; (X, \mathsf{d}))$, we notice that by construction one has $(\mathsf{e}_t)_{\#}\boldsymbol{\eta} = \mu_t$ for all $t \in D$. On the other hand, $\mu_t \in AC^2([0, T]; (\mathscr{P}(X), W_{\mathsf{d}}))$ implies, thanks to Proposition 5.2, that $t \mapsto \mu_t$ is continuous w.r.t. $\tau_{\mathscr{P}}$. Since d-convergence implies τ-convergence also $t \mapsto (\mathsf{e}_t)_{\#}\boldsymbol{\eta}$ is continuous w.r.t. $\tau_{\mathscr{P}}$, therefore $(\mathsf{e}_t)_{\#}\boldsymbol{\eta} = \mu_t$ for all $t \in [0, 1]$. An analogous approximation argument gives

$$\int \mathscr{A}_2(\eta, [s, t]) \, \mathrm{d}\boldsymbol{\eta}(\eta) \le \int_s^t |\dot{\mu}_r|^2 \, \mathrm{d}r \qquad \forall s, t \in [0, 1], \ s \le t \tag{5.13}$$

starting from (5.11).

Finally, from (5.13) and Fubini's theorem we get

$$\int |\dot{\eta}(t)|^2 \, \mathrm{d}\boldsymbol{\eta}(\eta) \le |\dot{\mu}_t|^2 \qquad \text{for a.e. } t \in (0, 1) . \tag{5.14}$$

On the other hand, since $(\mathsf{e}_s, \mathsf{e}_t)_{\#}\boldsymbol{\eta} \in \Gamma(\mu_s, \mu_t)$ one has

$$W_{\mathsf{d}}^2(\mu_s, \mu_t) \le \int \mathsf{d}^2(\eta(s), \eta(t)) \, \mathrm{d}\boldsymbol{\eta}(\eta) \le (t - s) \int \int_s^t |\dot{\eta}|^2(r) \, \mathrm{d}r \, \mathrm{d}\boldsymbol{\eta}(\eta) ,$$

hence for a.e. $t \in (0,1)$ the converse inequality to (5.14) holds. □

6. Cheeger energy and minimal relaxed slope

Throughout this section $(X, \tau, \mathsf{d}, \mathfrak{m})$ is an extended metric measure spaces according to Definition 4.7. In this section we provide basic calculus results already developed in [3], with minor variants in the definitions that do not really affect the proofs.

For $f \in \mathrm{Lip}_b(X, \tau, \mathsf{d})$, the asymptotic Lipschitz constant $\mathrm{Lip}_a(f, x) : X \to [0, \infty]$ is defined by

$$\mathrm{Lip}_a(f, x) = \lim_{r \downarrow 0} \mathrm{Lip}_a(f, x, r) \quad \text{with} \quad \mathrm{Lip}_a(f, x, r) := \sup_{\mathsf{d}(y,x) \vee \mathsf{d}(z,x) < r, \, \mathsf{d}(y,z) > 0} \frac{|f(y) - f(z)|}{\mathsf{d}(y, z)}, \quad (6.1)$$

and with the usual convention $\mathrm{Lip}_a(f, x) = 0$ at d-isolated points x. By construction the function $\mathrm{Lip}_a(f, \cdot)$ is d-upper semicontinuous. In the standard case when d is a finite distance and τ is the metric topology it follows that $\mathrm{Lip}_a(f, \cdot)$ is also τ-upper semicontinuous.

Definition 6.1 (Cheeger energy). *For all $f \in L^2(X, \mathfrak{m})$ we set*

$$\mathrm{Ch}(f) := \inf \liminf_{n \to \infty} \int g_n^2 \, \mathrm{d}\mathfrak{m}, \qquad D(\mathrm{Ch}) := \left\{ f \in L^2(X, \mathfrak{m}) : \ \mathrm{Ch}(f) < \infty \right\},$$

where the infimum runs among all sequences $(f_n) \subset \mathrm{Lip}_b(X, \tau, \mathsf{d})$ with $\lim_n \int |f_n - f|^2 \, \mathrm{d}\mathfrak{m} = 0$ and all $\mathfrak{m}$-measurable functions $g_n \geq \mathrm{Lip}_a(f_n, \cdot)$ $\mathfrak{m}$-a.e. in X.

Motivated by the previous definition we may define, for $f \in \mathrm{Lip}(X, \tau, \mathsf{d})$, $\mathrm{Lip}_a^*(f, \cdot)$ as the (essential) least upper bound of all $\mathfrak{m}$-measurable functions larger $\mathfrak{m}$-a.e. than $\mathrm{Lip}_a(f, \cdot)$. Then, Ch we can be equivalently defined by minimizing $\liminf_n \int (\mathrm{Lip}_a^*(f_n, \cdot))^2 \, \mathrm{d}\mathfrak{m}$ among all sequences $(f_n) \subset \mathrm{Lip}_b(X, \tau, \mathsf{d})$ with $\lim_n \int |f_n - f|^2 \, \mathrm{d}\mathfrak{m} = 0$.

The concept of minimal relaxed slope is closely related to the definition of Ch. First, one defines *relaxed slope* of f any function $G \geq g$, with g weak $L^2(X, \mathfrak{m})$ limit point as $n \to \infty$ of $\mathrm{Lip}_a^*(f_n, \cdot)$, where $f_n \in \mathrm{Lip}_b(X, \tau, \mathsf{d})$ and $f_n \to f$ in $L^2(X, \mathfrak{m})$. It can be proved ([3, Lem. 4.3]) that the class of relaxed slopes is a convex closed subset of $L^2(X, \mathfrak{m})$, not empty if and only if $f \in D(\mathrm{Ch})$. The minimal relaxed slope, denoted $|\mathrm{D}f|_w$ (and occasionally by $|\mathrm{D}f|_{w,\mathsf{d}}$ to emphasize its dependence on d), is the relaxed slope with smallest $L^2(X, \mathfrak{m})$ norm.

In analogy with the classical case, for all $f \in L^2(X, \mathfrak{m})$ with $\partial \mathrm{Ch}(f) \neq \emptyset$ we denote by Δf the element with minimal $L^2(X, \mathfrak{m})$ norm in $\partial \frac{1}{2} \mathrm{Ch}(f)$.

We now recall some basic calculus rules and more precise relations between Ch and the minimal relaxed slope. Properties (g) and (h) below involve the notion of test plan, recalled below.

Definition 6.2 (Test plan). *We say that $\boldsymbol{\eta} \in \mathscr{P}(X^{[0,1]})$ is a 2-test plan (relative to $\mathfrak{m}$) if $\boldsymbol{\eta}$ is concentrated on $AC^2([0,1]; (X, \mathsf{d}))$ and there exists $C \in [0, \infty)$ satisfying $(\mathrm{e}_t)_{\#} \boldsymbol{\eta} \leq C\mathfrak{m}$ for all $t \in [0,1]$. The least constant C with this property will be denoted by $C(\boldsymbol{\eta})$.*

Proposition 6.3. *The following properties hold:*

(a) *For all $f, g \in D(\mathrm{Ch})$, $\alpha, \beta \in \mathbb{R}$*

$$|\mathrm{D}(\alpha f + \beta g)|_w \leq |\alpha| \, |\mathrm{D}f|_w + |\beta| \, |\mathrm{D}g|_w \ ; \tag{6.2}$$

in particular Ch and $(\mathrm{Ch})^{1/2}$ are convex and lower semicontinuous functionals in $L^2(X, \mathfrak{m})$, with a dense domain.

(b) *For all $f \in D(\mathrm{Ch})$ one has $\mathrm{Ch}(f) = \int |\mathrm{D}f|_w^2 \, \mathrm{d}\mathfrak{m}$ and there exist $f_n \in \mathrm{Lip}_b(X, \tau, \mathsf{d})$ with $f_n \to f$ in $L^2(X, \mathfrak{m})$ and $\mathrm{Lip}_a^*(f_n, \cdot) \to |\mathrm{D}f|_w$ in $L^2(X, \mathfrak{m})$.*

(c) *$|\mathrm{D}f|_w = |\mathrm{D}g|_w$ $\mathfrak{m}$-a.e. in $\{f = g\}$ for all $f, g \in D(\mathrm{Ch})$.*

(d) *$|\mathrm{D}f|_w \leq \mathrm{Lip}_a^*(f, \cdot)$ $\mathfrak{m}$-a.e. in X for all $f \in \mathrm{Lip}_b(X, \tau, \mathsf{d})$.*

(e) *$|\mathrm{D}\phi(f)|_w = |\phi'(f)||\mathrm{D}f|_w$ $\mathfrak{m}$-a.e. in X, for all $f \in D(\mathrm{Ch})$ and $\phi : \mathbb{R} \to \mathbb{R}$ Lipschitz.*

(f) *$\int f \Delta g \, \mathrm{d}\mathfrak{m} \leq \int |\mathrm{D}f|_w |\mathrm{D}g|_w \, \mathrm{d}\mathfrak{m}$ for all $f \in D(\mathrm{Ch})$, $g \in D(\Delta)$.*

(g) *If $\boldsymbol{\eta} \in \mathscr{P}(X^{[0,1]})$ is a test plan, then for all $f \in D(\mathrm{Ch})$ one has*

$$|f(\eta(1)) - f(\eta(0))| \leq \int_0^1 |\mathrm{D}f|_w(\eta(s))|\dot{\eta}(s)| \, \mathrm{d}s \quad \text{for } \boldsymbol{\eta}\text{-a.e. } \eta \ .$$

(h) *If $\boldsymbol{\eta} \in \mathscr{P}(X^{[0,1]})$ is a test plan, then for all $f \in D(\mathrm{Ch})$ one has*

$$\limsup_{t \downarrow 0} \int \frac{|f(\eta(t)) - f(\eta(0))|^2}{(E_t(\eta))^2} \, \mathrm{d}\boldsymbol{\eta} \leq \int |\mathrm{D}f|_w^2(\eta(0)) \, \mathrm{d}\boldsymbol{\eta}(\eta) \,,$$

where $E_t(\eta) := \sqrt{t \int_0^t |\dot{\eta}(s)|^2 \, \mathrm{d}s}$.

Proof. The properties from (a) to (e) are proved in Lemma 4.3, Lemma 4.4, Theorem 4.5, and Proposition 4.8 of [3]. The proof of (f) relies on the convexity inequality $|\mathrm{D}(f + \varepsilon g)|_w \leq |\mathrm{D}f|_w + \varepsilon |\mathrm{D}g|$ with $\varepsilon > 0$, see Proposition 4.15 of [3]. Property (g) is proved in Theorem 5.14, Corollary 5.15 (see also Corollary 3.15) of [3], using Mazur's lemma, property (b) and the upper gradient property of the asymptotic Lipschitz constant. Finally, (h) follows by (g) and Hölder's inequality, which give

$$\int \frac{|f(\eta(t)) - f(\eta(0))|^2}{(E_t(\eta))^2} \, \mathrm{d}\boldsymbol{\eta} \leq \frac{1}{t} \int_0^t \int |\mathrm{D}f|_w^2 \, \mathrm{d}(\mathrm{e}_s)_\# \boldsymbol{\eta} \, \mathrm{d}s \,.$$

$\square$

Corollary 6.4. *If (X, d) is complete, for all $g \in D(\mathrm{Ch})$ and all $\mu_t = \rho_t \mathfrak{m} \in AC^2([0,T]; (\mathscr{P}(X), W_\mathsf{d}))$ with $\sup_t \|\rho_t\|_{L^\infty(X, \mathfrak{m})} < \infty$ one has*

$$\left| \int g \rho_T \, \mathrm{d}\mathfrak{m} - \int g \rho_0 \, \mathrm{d}\mathfrak{m} \right| \leq \int_0^T \left(\int |\mathrm{D}g|_w^2 \rho_t \, \mathrm{d}\mathfrak{m} \right)^{1/2} |\dot{\mu}_t| \, \mathrm{d}t \,.$$

Proof. By Proposition 5.5 we can find $\boldsymbol{\eta} \in \mathscr{P}(X^{[0,T]})$ concentrated on $AC^2([0,T]; (X, \mathsf{d}))$ with $(\mathrm{e}_t)_\# \boldsymbol{\eta} = \rho_t \mathfrak{m}$ for all $t \in [0,T]$ and satisfying (5.10), so that Proposition 6.3(g) gives

$$g(\eta(T)) \leq g(\eta(0)) + \int_0^T |\mathrm{D}f|_w(\eta(t)) |\dot{\eta}(t)| \, \mathrm{d}t \quad \text{for } \boldsymbol{\eta}\text{-a.e. } \eta$$

By integrating this inequality and using Hölder's inequality with t fixed together with (5.10) the proof is achieved. $\square$

In the sequel we denote by

$$\mathsf{F}(\rho) := 4 \, \mathrm{Ch}(\sqrt{\rho}) \,, \qquad \rho \in L^1_+(X, \mathfrak{m})$$

the so-called Fisher information functional. Let us recall its main properties (see [3] for the simple proof).

Proposition 6.5. F *is a convex and L^1-lower semicontinuous functional in $L^1_+(X, \mathfrak{m})$. If $\sqrt{\rho} \in D(\mathrm{Ch})$ we have the equivalent representation*

$$\mathsf{F}(\rho) = \int_{\{\rho > 0\}} \frac{|\mathrm{D}\rho|_w^2}{\rho} \, \mathrm{d}\mathfrak{m} \,.$$

We denote by P_t the $L^2(X, \mathfrak{m})$ (metric) gradient flow of the convex and lower semicontinuous functional $\frac{1}{2}\mathrm{Ch}$. Since $D(\mathrm{Ch})$ includes $\mathrm{Lip}_b(X, \tau, \mathsf{d})$ which is dense in $L^2(X, \mathfrak{m})$, P_t is a contraction semigroup in $L^2(X, \mathfrak{m})$, characterized by

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathsf{P}_t f = \Delta \mathsf{P}_t f \qquad \text{for a.e. } t > 0 \,. \tag{6.3}$$

Besides the general properties of gradient flows in Hilbert spaces listed in Proposition 3.4, we recall that P_t satisfies the maximum and minimum principle (i.e. if $c \leq f \leq C$ $\mathfrak{m}$-a.e. in X, then $c \leq \mathsf{P}_t f \leq C$ $\mathfrak{m}$-a.e. in X for all $t \geq 0$) and that

$$\left| \frac{\mathrm{d}}{\mathrm{d}t} \int f \mathsf{P}_t \rho \, \mathrm{d}\mathfrak{m} \right| \leq \mathsf{F}^{1/2}(\mathsf{P}_t \rho) \left(\int |\mathrm{D}f|_w^2 \mathsf{P}_t \rho \, \mathrm{d}\mathfrak{m} \right)^{1/2} \qquad \text{for a.e. } t \in (0, \infty) \,. \tag{6.4}$$

The maximum and minimum principle can be derived from Proposition 6.3(e), while (6.4) is a direct consequence of Proposition 6.3(f) and of (6.3).

Note also the following fact (whose proof can be obtained by a simple regularization argument, since $\mathfrak{m}$ is finite, see [3]): for all $f \in L^2_+(X, \mathfrak{m})$ the function $t \mapsto \int f_t \log f_t \, d\mathfrak{m}$ is absolutely continuous in $[0, \infty)$ and

$$\frac{d}{dt} \int f_t \log f_t \, d\mathfrak{m} = - \int_{\{f_t > 0\}} \frac{|Df_t|^2_w}{f_t} \, d\mathfrak{m} \qquad \text{for a.e. } t > 0 . \tag{6.5}$$

In particular, the right hand side is locally integrable in $[0, \infty)$.

7. EXTENDED DISTANCES IN $\mathscr{P}^a(X)$

In this section we introduce a class of absolutely continuous curves in an extended metric measure space $(X, \tau, d, \mathfrak{m})$, following the analogy with [2, Thm. 8.3.1], [9], [27].

We first introduce a Banach structure on two Sobolev classes of test functions, $D(\mathrm{Ch})$ and the algebra

$$\mathcal{A}_{\mathrm{Ch}} := \left\{ f \in D(\mathrm{Ch}) : \ f, |Df|_w \in L^\infty(X, \mathfrak{m}) \right\} \tag{7.1}$$

which obviously includes $\mathrm{Lip}_b(X, \tau, d)$ and it is dense in $L^2(X, \mathfrak{m})$; a simple truncation argument also shows that $\mathcal{A}_{\mathrm{Ch}}$ is dense in $L^p(X, \mathfrak{m})$ for every $p \in [1, \infty)$.

Since $f \mapsto \sqrt{\mathrm{Ch}(f)}$ is convex in $D(\mathrm{Ch})$ and the function $(x, y) \mapsto \sqrt{x^2 + y^2}$ is a norm in $\mathbb{R}^2$, it is easy to check that

$$\|f\|_{\mathrm{Ch}} := \left(\|f\|^2_2 + \mathrm{Ch}(f) \right)^{1/2} = \left(\int \left(|f|^2 + |Df|^2_w \right) d\mathfrak{m} \right)^{1/2} , \tag{7.2}$$

is a norm in $D(\mathrm{Ch})$; the lower semicontinuity of Ch with respect to L^2-convergence also shows that $\left(D(\mathrm{Ch}), \| \cdot \|_{D(\mathrm{Ch})} \right)$ is a Banach space. Similarly, $\mathcal{A}_{\mathrm{Ch}}$ is a Banach algebra w.r.t. the norm

$$\|f\|_{\mathcal{A}_{\mathrm{Ch}}} := \|f\|_\infty + \big\| |Df|_w \big\|_\infty . \tag{7.3}$$

$D(\mathrm{Ch})$ and the algebra $\mathcal{A}_{\mathrm{Ch}}$ are not separable in general, but since their norms are lower semicontinuous w.r.t. the L^2 convergence, they are F_σ and thus Borel subsets of $L^2(X, \mathfrak{m})$. When we will consider measurability of maps ϕ with values in $D(\mathrm{Ch})$ or $\mathcal{A}_{\mathrm{Ch}}$, we will always refer to its Borel σ-algebra inherited from the L^2 topology.

7.1. The dynamic approach and the continuity inequality.

Definition 7.1 (Continuity inequality). *Given a family of probability densities ρ_t, $t \in [0, T]$, we write $\rho_t \in \mathsf{CE}^2(X, \mathrm{Ch}, \mathfrak{m})$ if there exists $c \in L^2(0, T)$ satisfying*

$$\left| \int f \rho_t \, d\mathfrak{m} - \int f \rho_s \, d\mathfrak{m} \right| \leq \int_s^t c(r) \left(\int |Df|^2_w \rho_r \, d\mathfrak{m} \right)^{1/2} dr \tag{7.4}$$

for all $f \in \mathcal{A}_{\mathrm{Ch}}$ and all $0 \leq s \leq t \leq T$. The least c in (7.4) is denoted $\|\rho'_t\|$.

For simplicity of notation we do not emphasize the T dependence in the previous definition.

We will often deal with the case when ρ_t are essentially bounded, uniformly w.r.t. time: in this case, by using the density of $\mathcal{A}_{\mathrm{Ch}}$ in $L^1(X, \mathfrak{m})$, it is then easy to check that $t \mapsto \rho_t$ is weakly* continuous with values in $L^\infty(X, \mathfrak{m})$ (see also the next Theorem 7.3); we will write $\rho \in C_{\mathrm{w}*}([0, T]; L^\infty(X, \mathfrak{m}))$.

Remark 7.2. It is a direct consequence of (6.4) and (6.5) that $t \mapsto \rho_t := \mathsf{P}_t \rho$ belongs to $\mathsf{CE}^2(X, \mathrm{Ch}, \mathfrak{m})$ for all $\rho \in L^2_+(X, \mathfrak{m})$, with

$$\|\rho'_t\|^2 \leq \int_{\{\rho_t > 0\}} \frac{|D\rho_t|^2_w}{\rho_t} \, d\mathfrak{m} \qquad \text{for a.e. } t \in (0, \infty) .$$

In the following theorem we show a "differential" characterization of absolutely continuous curves in $(\mathscr{P}(X), W_d)$, which provides a key link between the metric and the differentiable viewpoints.

Theorem 7.3 (Differential characterization of absolutely continuous curves)**.**
For all $\rho_t \in \mathsf{CE}^2(X, \mathrm{Ch}, \mathfrak{m})$ one has $\mu_t = \rho_t \mathfrak{m} \in AC^2([0,T]; (\mathscr{P}(X), W_\mathsf{d}))$ and

$$|\dot{\mu}_t| \leq \|\rho_t'\| \qquad \text{for a.e. } t \in (0,T) \ . \tag{7.5}$$

Conversely, if (X, d) is complete, $\mu_t = \rho_t \mathfrak{m} \in AC^2([0,T]; (\mathscr{P}(X), W_\mathsf{d}))$ and $\sup_{t \in [0,T]} \|\rho_t\|_\infty < \infty$, then $\rho_t \in \mathsf{CE}^2(X, \mathrm{Ch}, \mathfrak{m})$ and

$$\|\rho_t'\| \leq |\dot{\mu}_t| \qquad \text{for a.e. } t \in (0,T) \ . \tag{7.6}$$

Proof. We show the second part of the statement and (7.6). Let $\mu_t = \rho_t \mathfrak{m} \in AC^2([0,T]; (\mathscr{P}(X), W_\mathsf{d}))$ with essentially bounded densities ρ_t.

The inequality (5.7) shows that $t \mapsto \int f \, \mathrm{d}\mu_t$ is absolutely continuous in $[0,T]$ for all $f \in \mathrm{Lip}_b(X, \tau, \mathsf{d})$. In addition, Corollary 6.4 provides the inequality

$$\left| \int f \rho_s \, \mathrm{d}\mathfrak{m} - \int f \rho_t \, \mathrm{d}\mathfrak{m} \right| \leq \int_s^t |\dot{\mu}_r| \left(\int |\mathrm{D}f|_w^2 \rho_r \, \mathrm{d}\mathfrak{m} \right)^{1/2} \mathrm{d}r$$

for $0 \leq s \leq t \leq T$. By the density of $\mathrm{Lip}_b(X, \tau, \mathsf{d})$ in $\mathcal{A}_{\mathrm{Ch}}$ provided by Proposition 6.3(b) the inequality extends to all $f \in \mathcal{A}_{\mathrm{Ch}}$.

We provide a proof of the converse implication and the converse inequality (7.5), along the lines of [32], assuming for simplicity $T = 1$. First we notice that the property $\rho_t \in \mathsf{CE}(X, \mathrm{Ch}, \mathfrak{m})$ is stable under convolution w.r.t. the time parameter: more precisely, if we extend $t \mapsto \rho_t$ by continuity and with constant values to $(-\infty, 0) \cup (1, \infty)$, then $\rho_{\varepsilon,t} := \rho_t * \chi_\varepsilon$ still belongs to $\mathsf{CE}(X, \mathrm{Ch}, \mathfrak{m})$ and $\|\rho_{\varepsilon,t}'\|^2 \leq \|\rho_t'\|^2 * \chi_\varepsilon$. For this reason, in the proof of this implication we can assume with no loss of generality that $t \mapsto \rho_t$ is continuous w.r.t. the $L^1(X, \mathfrak{m})$ topology. We start from the duality formula (5.2). Let $\phi : X \to [0, \infty)$ be such that $\phi \in C(K)$ and $\phi \equiv \max_K \phi$ on $X \setminus K$, with $K \subset X$ compact not empty. Under this restriction on ϕ, we have already seen that $Q_\varepsilon \phi$ can be represented in the form (5.4), and that $Q_\varepsilon \phi$ is d-Lipschitz, Borel (because it is τ-lower semicontinuous), nonnegative and bounded. In addition $Q_\varepsilon \phi \uparrow \phi$ and $Q_1(Q_\varepsilon \phi) \uparrow Q_1 \phi$ as $\varepsilon \downarrow 0$.

Set now $\varphi := Q_\varepsilon \phi$ for some $\varepsilon > 0$ and observe that $Q_t \varphi$, $t \in [0, 1]$, are uniformly d-Lipschitz and that the map $t \mapsto Q_t \varphi$ is Lipschitz from $[0, 1]$ with values in $L^\infty(X, \mathfrak{m})$. By applying [2, Lem. 4.3.4] to the function $(s, t) \mapsto \int \rho_s Q_t \varphi \, \mathrm{d}\mathfrak{m}$ we obtain that $t \mapsto \int \rho_t Q_t \varphi$ is absolutely continuous in $[0, 1]$ and that its derivative can be estimated from above by

$$\limsup_{s \to t} \frac{1}{|s - t|} \left| \int (\rho_s - \rho_t) Q_t \varphi \, \mathrm{d}\mathfrak{m} \right| + \limsup_{s \to t} \frac{1}{|s - t|} \left| \int \rho_t (Q_s \varphi - Q_t \varphi) \, \mathrm{d}\mathfrak{m} \right| .$$

Using the inequality

$$\left| \int \rho_t Q_t \varphi \, \mathrm{d}\mathfrak{m} - \int \rho_s Q_t \varphi \, \mathrm{d}\mathfrak{m} \right| \leq \int_s^t \|\rho_r'\| \left(\int |\mathrm{D}Q_t \varphi|_w^2 \rho_r \, \mathrm{d}\mathfrak{m} \right)^{1/2} \mathrm{d}r$$

we estimate the first limsup, at Lebesgue points t of $s \mapsto |\rho_s'|$, by

$$\|\rho_t'\| \left(\int |\mathrm{D}Q_t \varphi|_w^2 \rho_t \, \mathrm{d}\mathfrak{m} \right)^{1/2}$$

(here we used also the strong continuity of $s \mapsto \rho_s$). Estimating the second lim sup with Fatou's lemma and using Proposition 6.3(d) gives

$$\int (\rho_1 Q_1 \varphi - \rho_0 \varphi) \, \mathrm{d}\mathfrak{m} \leq \int_0^1 \left[\|\rho_t'\| \left(\int (\mathrm{Lip}_a^*(Q_t \varphi, \cdot))^2 \rho_t \, \mathrm{d}\mathfrak{m} \right)^{1/2} + \int \rho_t \xi_t \, \mathrm{d}\mathfrak{m} \right] \mathrm{d}t ,$$

where ξ_t is the bounded Borel function

$$\xi_t := \limsup_{s \to t} \frac{Q_s \varphi - Q_t \varphi}{s - t} \ .$$

Now we use the pointwise subsolution property

$$\frac{1}{2} \left(\mathrm{Lip}_a^*(Q_t \varphi, \cdot) \right)^2 \leq -\xi_t \qquad \mathfrak{m}\text{-a.e. in } X \tag{7.7}$$

for a.e. $t \in (0,1)$ (whose proof follows as in [3, Thm. 3.5], where it is stated in a weaker form with the slope in place of the asymptotic Lipschitz constant, see also [6]) and the Young inequality to get $\int (\rho_1 Q_1 \varphi - \rho_0 \varphi) \, \mathrm{d}\mathfrak{m} \leq \frac{1}{2} \int_0^1 \|\rho_t'\|^2 \, \mathrm{d}t$. Remembering that $\varphi = Q_\varepsilon \phi$, we can let $\varepsilon \downarrow 0$ and use the arbitrariness of ϕ to get

$$W_{\mathsf{d}}^2(\rho_1 \mathfrak{m}, \rho_0 \mathfrak{m}) \leq \int_0^1 \|\rho_t'\|^2 \, \mathrm{d}t \ .$$

By applying this inequality to a rescaled version of ρ we obtain $W_{\mathsf{d}}^2(\rho_t \mathfrak{m}, \rho_s \mathfrak{m}) \leq (s-t) \int_t^s \|\rho_r'\|^2 \, \mathrm{d}r$ for all $s, t \in [0,1]$ with $s > t$, so that by differentiation the inequality (7.5) follows at a.e. $t \in (0,1)$. $\qquad \square$

Using the continuity inequality we can define an extended "Wasserstein-like" distance W_{Ch} in $\mathscr{P}^a(X)$ in the same spirit of the Benamou-Brenier formula:

$$W_{\mathrm{Ch}}^2(\rho_0 \mathfrak{m}, \rho_1 \mathfrak{m}) := \inf \left\{ \int_0^1 \|\rho_t'\|^2 \, \mathrm{d}t : \ \rho_t \in \mathsf{CE}^2(X, \mathrm{Ch}, \mathfrak{m}) \right\} . \tag{7.8}$$

This definition is also natural in view of Remark 7.2. Even though it is conceptually convenient to think to W_{Ch} as an extended distance in $\mathscr{P}^a(X)$, we occasionally adopt we the simpler notation $W_{\mathrm{Ch}}(\rho_0, \rho_1)$, i.e. we identify measures in $\mathscr{P}^a(X)$ with their densities w.r.t. $\mathfrak{m}$. The same remark applies to the other distances in $\mathscr{P}^a(X)$ we shall introduce. Now we provide a few basic properties of W_{Ch}.

Proposition 7.4 (Properties of W_{Ch}).
  (a) $(\mathscr{P}^a(X), W_{\mathrm{Ch}})$ *is an extended length metric space, and* $W_{\mathrm{Ch}} \geq W_{\mathsf{d}}$.
  (b) *Assume that* $\mu_t^n = \rho_t^n \mathfrak{m} \in \mathscr{P}(X)$ *satisfy* $\rho_t^n \to \rho_t$ *weakly in* $L^1(X, \mathfrak{m})$ *for all* $t \in [0, T]$ *and that* $\rho^n \in \mathsf{CE}^2(X, \mathrm{Ch}, \mathfrak{m})$ *with* $\|(\rho^n)_t'\|$ *uniformly bounded in* $L^2(0, T)$. *Then, if* $c(t)$ *is any weak limit point of* $\|(\rho^n)_t'\|$ *as* $n \to \infty$, *one has* $\rho_t \in \mathsf{CE}^2(X, \mathrm{Ch}, \mathfrak{m})$ *with* $\|\rho_t'\| \leq c(t)$ *for a.e.* $t \in (0, T)$.
  (c) W_{Ch}^2 *is jointly convex in* $(\mathscr{P}^a(X))^2$.

Proof. From (7.5) with $T = 1$ we obtain that $W_{\mathsf{d}}(\mu, \nu) \leq W_{\mathrm{Ch}}(\mu, \nu)$ whenever $\mu, \nu \in \mathscr{P}^a(X)$. This yields immediately that $W_{\mathrm{Ch}}(\mu, \nu) = 0$ implies $\mu = \nu$. The proof of the triangle property of W_{Ch} follows by a standard concatenation argument, noticing that for any $T > 0$ one has

$$W_{\mathrm{Ch}}(\rho_0, \rho_T) := \inf \left\{ \int_0^T \|\rho_t'\| \, \mathrm{d}t : \ \rho_t \in \mathsf{CE}^2(X, \mathrm{Ch}, \mathfrak{m}) \right\} . \tag{7.9}$$

The length property also follows directly from (7.9), while the proof of (b) is a direct consequence of a passage to the limit as $n \to \infty$ in (7.4).

In order to prove (c), notice that a convex combination of (7.4) written for $\rho_t, \sigma_t \in \mathsf{CE}^2(X, \mathrm{Ch}, \mathfrak{m})$ gives

$$\left| \int f((1-\alpha)\rho_t + \alpha \sigma_t) \, \mathrm{d}\mathfrak{m} - \int f((1-\alpha)\rho_s + \alpha \sigma_s) \, \mathrm{d}\mathfrak{m} \right|$$
$$\leq \int_s^t (1-\alpha)\|\rho_r'\| \left(\int |\mathrm{D}f|_w^2 \rho_r \, \mathrm{d}\mathfrak{m} \right)^{1/2} + \alpha \|\sigma_r'\| \left(\int |\mathrm{D}f|_w^2 \sigma_r \, \mathrm{d}\mathfrak{m} \right)^{1/2} \mathrm{d}r$$
$$\leq \int_s^t \sqrt{(1-\alpha)\|\rho_r'\|^2 + \alpha \|\sigma_r'\|^2} \left(\int |\mathrm{D}f|_w^2 ((1-\alpha)\rho_r + \alpha \sigma_r) \, \mathrm{d}\mathfrak{m} \right)^{1/2} \mathrm{d}r$$

for $0 \leq s \leq t \leq T$. It follows that

$$W_{\mathrm{Ch}}^2\big((1-\alpha)\rho_1 + \alpha \sigma_1, (1-\alpha)\rho_0 + \alpha \sigma_0 \big) \leq \int_0^1 (1-\alpha)\|\rho_r'\|^2 + \alpha \|\sigma_r'\|^2 \, \mathrm{d}r$$

and, by minimizing, we conclude. $\qquad \square$

In the following corollary we reverse the inequality $W_{\mathsf{d}} \leq W_{\mathrm{Ch}}$ on probability measures with density in $L^\infty(X, \mathfrak{m})$, at the level of absolutely continuous curves and metric derivatives.

Corollary 7.5 (Equality of metric derivatives)**.** *Assume that (X, d) is complete and let $(\rho_t)_{t \in [0,T]}$ be a curve of probability densities with $\sup_{t \in [0,T]} \|\rho_t\|_\infty < \infty$. Then, for $\mu_t = \rho_t \mathbf{m}$, one has*

$$\mu_t \in AC^2([0,T]; (\mathscr{P}(X), W_\mathsf{d})) \qquad \Longleftrightarrow \qquad \mu_t \in AC^2([0,T]; (\mathscr{P}(X), W_\mathrm{Ch}))$$

and the corresponding metric derivatives coincide a.e. in $(0, T)$.

Proof. The implication $\Leftarrow$ is obvious, because $W_\mathrm{Ch} \geq W_\mathsf{d}$. In order to prove the converse one, first apply the first part of the statement of Theorem 7.3 to obtain $\rho_t \in \mathsf{CE}^2(X, \mathrm{Ch}, \mathbf{m})$ and $\|\rho_t'\| \leq |\mu_t'| \in L^2(0, T)$. By the very definition of W_Ch, this implies $\mu_t \in AC^2([0,T]; (\mathscr{P}(X), W_\mathrm{Ch}))$. The coincidence of the metric derivatives is a simple consequence of (7.5), (7.6). $\qquad \square$

7.2. **A dual distance induced by subsolutions of the Hamilton-Jacobi equation.** We close this section by introducing another "dual" extended distance $W_{\mathrm{Ch},*}$ in $\mathscr{P}^a(X)$, motivated by the analogy with the dual formulation of the optimal transport problem, the inequality $Q_1 f(x) - f(y) \leq \frac{1}{2} \mathsf{d}^2(x, y)$ and the subsolution property (7.7) of $Q_t f$.

Definition 7.6. *For $\mu_0 = \rho_0 \mathbf{m}$, $\mu_1 = \rho_1 \mathbf{m} \in \mathscr{P}^a(X)$ we define*

$$W_{\mathrm{Ch},*}^2(\rho_0, \rho_1) := 2 \sup_\phi \int (\phi_1 \rho_1 - \phi_0 \rho_0) \, \mathrm{d}\mathbf{m} \,, \tag{7.10}$$

where the supremum runs in the convex set of all the bounded Borel maps $\phi(t, x) = \phi_t(x)$ satisfying $\phi \in C_{\mathrm{w}*}([0,1]; L^\infty(X, \mathbf{m})) \cap L^1(0, 1; D(\mathrm{Ch}))$, and

$$\frac{\mathrm{d}}{\mathrm{d}t} \phi_t + \frac{1}{2} |\mathrm{D}\phi_t|_w^2 \leq 0 \qquad \text{in } (0,1) \times X, \text{ in the duality with } \mathcal{A}_\mathrm{Ch} \,. \tag{7.11}$$

The inequality (7.11) has to be understood as

$$\frac{\mathrm{d}}{\mathrm{d}t} \int \phi_t \psi \, \mathrm{d}\mathbf{m} + \frac{1}{2} \int \psi |\mathrm{D}\phi_t|_w^2 \, \mathrm{d}\mathbf{m} \leq 0 \qquad \text{in } \mathcal{D}'(0,1) \tag{7.12}$$

for all $\psi \in \mathcal{A}_\mathrm{Ch}$ nonnegative.

Lemma 7.7 (Equivalent admissible class of subsolutions to (7.11))**.** *The supremum in formula (7.10) can be equivalently taken w.r.t. subsolutions ϕ to (7.11) in the class $\phi \in C^\infty([0,1]; \mathcal{A}_\mathrm{Ch})$.*

Proof. By approximating any admissible ϕ in the definition of $W_{\mathrm{Ch},*}$ with the functions

$$\phi_\lambda(t, x) := \lambda \phi(\lambda t + (1 - \lambda)/2, x)$$

and by letting $\lambda \uparrow 1$, we see that is not restrictive to assume the existence of $a < 0$ and $b > 1$ such that ϕ is bounded, $\phi \in C_{\mathrm{w}*}([a,b]; L^\infty(X, \mathbf{m})) \cap L^1(a, b; D(\mathrm{Ch}))$ and $\partial_t \phi_t + |\mathrm{D}\phi_t|_w^2/2 \leq 0$ in $(a, b) \times X$ according to (7.12). Then, by mollification w.r.t. to t, which preserves the Hamilton-Jacobi subsolution property, we can also assume that $\phi \in C^\infty([0,1]; L^\infty(X, \mathbf{m})) \cap C^\infty([0,1]; D(\mathrm{Ch}))$ with $\phi(\cdot, x) \in C^k([0,1])$, uniformly w.r.t. x. Under this assumption, the subsolution property is satisfied $\mathbf{m}$-a.e. in X, for all t, which also shows that the map $t \mapsto |\mathrm{D}\phi_t|_w$ is also uniformly bounded in $L^\infty(X, \mathbf{m})$. It follows that ϕ is uniformly bounded with values in $\mathcal{A}_\mathrm{Ch}$ and strongly measurable with respect to the Borel sets induced by the L^2-topology. A further convolution in time (or the mollification by a semigroup in the first step) shows that we can also assume $\phi \in C^k([0,1]; \mathcal{A}_\mathrm{Ch})$. $\qquad \square$

Remark 7.8 (Elementary properties of $W_{\mathrm{Ch},*}$)**.**
(1) By the scaling argument $\hat{\phi}(t, x) = \delta \phi(\delta t, x)$, it is easily seen that

$$W_{\mathrm{Ch},*}^2(\rho_0, \rho_1) = 2 \sup_{(\delta, \phi)} \delta \int (\phi_\delta \rho_1 - \phi_0 \rho_0) \, \mathrm{d}\mathbf{m} \,,$$

where the supremum runs among all pairs (δ, ϕ) with $\delta > 0$ and ϕ bounded Borel map $\phi(t, x) = \phi_t(x)$ satisfying $\phi \in C_{\mathrm{w}*}([0,\delta]; L^\infty(X, \mathbf{m})) \cap L^1(0, \delta; D(\mathrm{Ch}))$, and

$$\frac{\mathrm{d}}{\mathrm{d}t} \phi_t + \frac{1}{2} |\mathrm{D}\phi_t|_w^2 \leq 0 \qquad \text{in } (0,\delta) \times X, \text{ in the duality with } \mathcal{A}_\mathrm{Ch} \,.$$

(2) More generally, suppose that $\varphi \in C_{w*}([a,b]; L^\infty(X, \mathbf{m})) \cap L^1(a, b; D(\mathrm{Ch}))$ satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi_t + \frac{\vartheta(t)}{2}|\mathrm{D}\varphi_t|_w^2 \le 0 \qquad \text{in } (a, b) \times X, \text{ in the duality with } \mathcal{A}_{\mathrm{Ch}} \qquad (7.13)$$

where $\vartheta \in C([a,b])$ is a positive function. Then

$$2\alpha(b)\int (\rho_1\varphi_b - \rho_0\varphi_a) \, \mathrm{d}\mathbf{m} \le W^2_{\mathrm{Ch},*}(\rho_0, \rho_1) \quad \text{where} \quad \alpha(t) := \int_a^t \vartheta(r) \, \mathrm{d}r \qquad (7.14)$$

In fact, setting

$$\beta(t) := \alpha(t)/\alpha(b), \ t \in [a, b], \qquad \gamma(s) := \beta^{-1}(s)$$

so that γ is an increasing diffeomorphism between $[0,1]$ and $[a,b]$ satisfying $\gamma'(s) = \alpha(b)/\vartheta(\gamma(s))$, the curve $\widetilde{\varphi}_s := \alpha(b)\varphi_{\gamma(s)}$ solves

$$\frac{\mathrm{d}}{\mathrm{d}s}\widetilde{\varphi}_s + \frac{1}{2}|\mathrm{D}\widetilde{\varphi}_s|_w^2 \le 0 \qquad \text{in } (0, 1) \times X, \text{ in the duality with } \mathcal{A}_{\mathrm{Ch}}$$

so that

$$2\alpha(b)\int (\rho_1\varphi_b - \rho_0\varphi_a) \, \mathrm{d}\mathbf{m} = 2\int (\rho_1\widetilde{\varphi}_1 - \rho_0\widetilde{\varphi}_0) \, \mathrm{d}\mathbf{m} \le W^2_{\mathcal{E},*}(\rho_0, \rho_1) \ .$$

(3) It is not hard to prove that $W_{\mathrm{Ch},*}$ *is an extended distance:* indeed, the non-degeneracy condition follows by the inequality $W_{\mathrm{Ch},*} \ge W_{\mathsf{d}}$, proved in the next proposition. The symmetry property follows easily by replacing $\phi(t, x)$ by $-\phi(\delta - t, x)$. In order to prove the triangle inequality, given probability densities ρ, σ, λ, and constants $\delta > 0$ and $\delta' \in (0, \delta)$ we write

$$2\delta \int (\phi_\delta\lambda - \phi_0\rho) \, \mathrm{d}\mathbf{m} = 2\delta \int (\phi_\delta\lambda - \phi_{\delta'}\sigma) \, \mathrm{d}\mathbf{m} + 2\delta \int (\phi_{\delta'}\sigma - \phi_0\rho) \, \mathrm{d}\mathbf{m}$$

$$\le \frac{2\delta}{2(\delta - \delta')}W^2_{\mathrm{Ch},*}(\lambda, \sigma) + \frac{2\delta}{2\delta'}W^2_{\mathrm{Ch},*}(\sigma, \rho) \ .$$

Now we minimize w.r.t. δ' and use the identity $\inf_{(0,1)} s^{-1}a^2 + (1-s)^{-1}b^2 = (a+b)^2$ to get

$$2\delta \int \phi_\delta\lambda - \phi_0\rho \, \mathrm{d}\mathbf{m} \le \left(W_{\mathrm{Ch},*}(\lambda, \sigma) + W_{\mathrm{Ch},*}(\sigma, \rho)\right)^2 \ .$$

By taking the supremum w.r.t. (δ, ϕ) we conclude.

(4) $W^2_{\mathrm{Ch},*}$ *is jointly convex in* $\mathscr{P}^a(X) \times \mathscr{P}^a(X)$ *and l.s.c. with respect to the weak topology of* $L^1(X, \mathbf{m})$, since it is defined as a supremum of a family of continuous linear functionals on $L^1(X, \mathbf{m})$. In particular, every closed sublevel of the Entropy functional (2.12) in $\mathscr{P}^a(X)$ is complete with respect to $W_{\mathrm{Ch},*}$.

We can now refine the inequality between W_{Ch} and W_{d}, proving that $W_{\mathrm{Ch},*}$ is intermediate.

Proposition 7.9 (Comparison of W_{Ch}, $W_{\mathrm{Ch},*}$ and W_{d}). $W_{\mathsf{d}} \le W_{\mathrm{Ch},*} \le W_{\mathrm{Ch}}$ *on* $(\mathscr{P}^a(X))^2$.

Proof. We first prove that $W_{\mathrm{Ch}} \ge W_{\mathrm{Ch},*}$. By Lemma 7.7 we can assume that $\phi \in C^1([0, 1]; \mathcal{A}_{\mathrm{Ch}})$ with $\phi(\cdot, x) \in C^1([0, 1])$, uniformly w.r.t. x. Under this assumption, the subsolution property is satisfied $\mathbf{m}$-a.e. in X, for all t; in addition, for all $\rho_t \in \mathsf{CE}^2(X, \mathrm{Ch}, \mathbf{m})$, the Leibniz rule and a density argument easily give that $t \mapsto \int \phi_t\rho_t \, \mathrm{d}\mathbf{m}$ is absolutely continuous in $[0, 1]$, and that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int \phi_t\rho_t \, \mathrm{d}\mathbf{m} = \int \phi_t\frac{\mathrm{d}}{\mathrm{d}t}\rho_t \, \mathrm{d}\mathbf{m} + \int \rho_t\frac{\mathrm{d}}{\mathrm{d}t}\phi_t \, \mathrm{d}\mathbf{m} \le \int \phi_t\frac{\mathrm{d}}{\mathrm{d}t}\rho_t \, \mathrm{d}\mathbf{m} - \frac{1}{2}\int \rho_t|\mathrm{D}\phi_t|_w^2 \, \mathrm{d}\mathbf{m}$$

for a.e. $t \in (0, 1)$. By the Young inequality, it follows that

$$\left|\frac{\mathrm{d}}{\mathrm{d}t}\int \phi_t\rho_t \, \mathrm{d}\mathbf{m}\right| \le \frac{1}{2}\|\rho'_t\|^2 \qquad \text{for a.e. } t \in (0, 1) \ .$$

By integration in $(0, 1)$ and by minimizing w.r.t. ρ_t the inequality follows.

Now we prove that $W_{\mathrm{Ch},*} \ge W_{\mathsf{d}}$. Let (d_i) be an increasing net of bounded and $(\tau \times \tau)$-continuous semidistances with $\mathsf{d}_i \uparrow \mathsf{d}$. Taking Theorem 5.1 into account, it is sufficient to fix i

and prove that $W_{\mathsf{d}_i} \leq W_{\mathrm{Ch},*}$. In order to prove this property, taking (5.3) and the comments immediately after into account, it suffices to show that

$$Q_t^i \phi(x) := \inf_{y \in X} \phi(y) + \frac{1}{2t} \mathsf{d}_i^2(x, y)$$

is admissible in (7.10) whenever ϕ is bounded and d_i-Lipschitz (thus τ-continuous). This follows combining the subsolution property (see (7.7) and the comments after it)

$$\limsup_{s \to t} \frac{Q_s^i \phi(x) - Q_t^i \phi(x)}{s - t} + \frac{1}{2}\big(\mathrm{Lip}_{a,\mathsf{d}_i}^*(Q_t^i \phi, x)\big)^2 \leq 0 \quad \mathfrak{m}\text{-a.e. in } X$$

satisfied by $Q_t^i \phi$ for a.e. $t > 0$ with the inequalities

$$|\mathrm{D}Q_t^i \phi|_w(x) \leq \mathrm{Lip}_{a,\mathsf{d}}^*(Q_t^i \phi, x) \leq \mathrm{Lip}_{a,\mathsf{d}_i}^*(Q_t^i \phi, x) \qquad \mathfrak{m}\text{-a.e. in } X .$$

$\square$

Remark 7.10. One can also introduce the "dual" L^1 transport distance $W_{\mathrm{Ch},*,1}$:

$$W_{\mathrm{Ch},*,1}(\rho_0, \rho_1) := \sup_\phi \int \phi(\rho_1 - \rho_0) \, \mathrm{d}\mathfrak{m} , \tag{7.15}$$

where the supremum runs over all bounded and Borel maps $\phi \in D(\mathrm{Ch})$ with $|\mathrm{D}\phi|_w \leq 1$ $\mathfrak{m}$-a.e. in X. It is not hard to see that

$$W_{\mathrm{Ch},*,1}(\rho_0, \rho_1) \leq W_{\mathrm{Ch},*}(\rho_0, \rho_1) .$$

Indeed, fix ϕ with $|\mathrm{D}\phi|_w \leq 1$ and put $\phi_t(x) = -\frac{1}{2}t + \phi(x)$, which is admissible in the definition of $W_{\mathrm{Ch},*}$. Now for $\delta > 0$ we have

$$\int \phi(\rho_1 - \rho_0) \, \mathrm{d}\mathfrak{m} = \int (\phi_\delta \rho_1 - \phi_0 \rho_0) \, \mathrm{d}\mathfrak{m} + \frac{\delta}{2} \leq \frac{1}{2\delta} W_{\mathrm{Ch},*}^2(\rho_0, \rho_1) + \frac{\delta}{2} .$$

Optimizing in δ we find $\int \phi(\rho_1 - \rho_0) \, \mathrm{d}\mathfrak{m} \leq W_{\mathrm{Ch},*}(\rho_0, \rho_1)$ and taking the supremum over ϕ yields the claim.

8. IDENTIFICATION OF GRADIENT FLOWS

In this section we compare the metric gradient flows of Ent w.r.t. to the extended distances W_{d} and W_{Ch}, relating also them to the $L^2(X, \mathfrak{m})$ gradient flow P_t of $\frac{1}{2}\mathrm{Ch}$.

The following result is a small improvement of [3, Thm. 7.4], since we replace the slope of Ent w.r.t. W_{d} with the slope w.r.t. the (a priori larger) distance W_{Ch}. It can be obtained with the same proof.

Lemma 8.1 (The Fisher information is dominated by the slope of the Entropy)**.** *For every probability density* $\rho \in L_+^2(X, \mathfrak{m})$ *one has*

$$4 \int |\mathrm{D}\sqrt{\rho}|_w^2 \, \mathrm{d}\mathfrak{m} \leq |\mathrm{D}_{W_{\mathrm{Ch}}}^- \mathrm{Ent}|^2(\rho \, \mathfrak{m}) . \tag{8.1}$$

Proof. Let $\rho_t = \mathsf{P}_t \rho$; we set $\mu_t := \rho_t \mathfrak{m}$ and $\mu = \rho \mathfrak{m}$. Denoting by $|\dot\mu_t|$ the metric derivative w.r.t. W_{Ch}, from Remark 7.2 we get

$$|\dot\mu_t|^2 \leq \mathsf{F}(\rho_t) \qquad \text{for a.e. } t > 0 . \tag{8.2}$$

Applying (6.5) we get

$$\mathrm{Ent}(\mu) - \mathrm{Ent}(\mu_t) = \int_0^t \mathsf{F}(\rho_s) \, \mathrm{d}s \geq \frac{1}{2} \int_0^t \mathsf{F}(\rho_s) \, \mathrm{d}s + \frac{1}{2} \int_0^t |\dot\mu_s|^2 \, \mathrm{d}s \tag{8.3}$$

$$\geq \frac{1}{2}\Big(\frac{1}{\sqrt{t}} \int_0^t \sqrt{\mathsf{F}(\rho_s)} \, \mathrm{d}s\Big)^2 + \frac{1}{2}\Big(\frac{1}{\sqrt{t}} \int_0^t |\dot\mu_s| \, \mathrm{d}s\Big)^2 \geq \frac{1}{t}\Big(\int_0^t \sqrt{\mathsf{F}(\rho_s)} \, \mathrm{d}s\Big) W_{\mathrm{Ch}}(\mu, \mu_t) .$$

Dividing by $W_{\mathrm{Ch}}(\mu, \mu_t)$ and passing to the limit as $t \downarrow 0$ we get (8.1), since the lower semicontinuity of Ch yields

$$\sqrt{\mathsf{F}(\rho)} \leq \liminf_{t \downarrow 0} \frac{1}{t} \int_0^t \sqrt{\mathsf{F}(\rho_s)} \, \mathrm{d}s .$$

$\square$

In order to identify the metric gradient flows of Ent with P_t, we will also use the following result, see [3, Lem. 5.17, Thm. 8.1]. Its proof uses Proposition 5.5, Proposition 6.3 (e), (g), the estimates (6.4), (6.5), the convexity of F and the strict convexity of Ent, see also the next section for an analogous argument involving the same ingredients.

Theorem 8.2. *Let $(\rho_t)_{t \in [0,T]}$ be a curve of bounded probability densities with $\sup_t \|\rho_t\|_\infty < \infty$. Assume that $\mu_t = \rho_t \mathfrak{m} \in AC^2([0,T]; (\mathscr{P}(X), W_{\mathsf{d}}))$ and that μ_t satisfies the Entropy-Fisher dissipation inequality*

$$\operatorname{Ent}(\mu_0) \geq \operatorname{Ent}(\mu_T) + \frac{1}{2} \int_0^T |\dot\mu_t|^2 \, \mathrm{d}t + \frac{1}{2} \int_0^T \mathsf{F}(\rho_t) \, \mathrm{d}t \ . \tag{8.4}$$

Then $\rho_t = \mathsf{P}_t \rho_0$ for all $t \in [0,T]$ and equality holds in (8.4).

Theorem 8.3 (Identification of gradient flows)**.** *Let $(X, \tau, \mathsf{d}, \mathfrak{m})$ be an extended metric measure space with (X, d) complete. Let $(\rho_t)_{t \in [0,\infty)}$ be a curve of probability densities with $\|\rho_t\|_\infty \in L^\infty_{\mathrm{loc}}([0,\infty))$ and set $\mu_t = \rho_t \mathfrak{m}$ and let us consider the properties*

  (a) *μ_t is a metric gradient curve of Ent relative to W_{d} starting from μ_0;*
  (b) *μ_t is a metric gradient curve of Ent relative to W_{Ch} starting from μ_0;*
  (c) *$\rho_t = \mathsf{P}_t \rho_0$ for all $t \in [0,\infty)$.*

Then $(a) \Rightarrow (b) \Rightarrow (c)$. If $|\mathrm{D}_{\mathsf{d}}^- \operatorname{Ent}|$ is lower semicontinuous in $L^1_+(X, \mathfrak{m})$, then $(c) \Rightarrow (a)$.

Proof. By Corollary 7.5 and the inequality $W_{\mathsf{d}} \leq W_{\mathrm{Ch}}$, which yields a converse inequality at the level of slopes, the metric gradient curves in (a) are contained in the metric gradient curves in (b). On the other hand, by (8.1) of Lemma 8.1, the metric gradient curves in (b) satisfy the Entropy-Fisher dissipation inequality and therefore, thanks to Theorem 8.2, satisfy (c). Finally, under the lower semicontinuity assumption on $|\mathrm{D}_{\mathsf{d}}^- \operatorname{Ent}|$, the identity

$$\operatorname{Ent}(\rho\mathfrak{m}) - \operatorname{Ent}(\mathsf{P}_t\rho\mathfrak{m}) = \int_0^t \mathsf{F}(P_s\rho) \, \mathrm{d}s$$

and the inequality $\mathsf{F}(\rho) \geq |\mathrm{D}_{\mathsf{d}}^- \operatorname{Ent}|^2(\rho\,\mathfrak{m})$ (see [3, Thm. 7.6]) show that the class (c) is contained in the class (a). $\square$

Remark 8.4. By comparison, the implications above can also be stated with the distance $W_{\mathrm{Ch},*}$. This is possible because, according to Proposition 7.9, $W_{\mathrm{Ch},*}$ is intermediate between W_{Ch} and W_{d}.

9. A STABILITY RESULT FOR CHEEGER'S ENERGIES

In this section we consider an extended metric-topological space (X, τ, d) and a monotone family of $(\tau \times \tau)$-continuous approximating semidistances $\mathsf{d}_i : X \times X \to [0,\infty)$ as in Definition 4.1. Given $\mathfrak{m} \in \mathscr{P}(X)$, our goal is to prove a convergence results for the corresponding Cheeger energies. Since in view of the applications we have in mind we want to cover also the case when d_i are semidistances, we have to adapt the construction of Section 6, thought for (extended) distances, to this slightly more general setting.

Let $(X_i, \widetilde{\mathsf{d}}_i)$ be the quotient metric space, with $\pi^i : X \to X_i$ the canonical projection. We choose in X_i the standard topology τ_i generated by the metric structure, so that $\operatorname{Lip}(X_i, \tau_i, \widetilde{\mathsf{d}}_i)$ is a standard metric-topological space and $\pi^i : X \to X_i$ is continuous (thanks to the $(\tau \times \tau)$-continuity of d_i). Thanks to the continuity of π^i we can also define $\mathfrak{m}_i = (\pi^i)_{\#}\mathfrak{m} \in \mathscr{P}(X_i)$, thus providing the structure of metric measure space to X_i.

The map $g \mapsto \pi^i_*(g) = g \circ \pi^i$ provides a linear isometry of $L^2(X_i, \mathfrak{m}_i)$ into $L^2(X, \mathfrak{m})$. Then, denoting by $\mathcal{D}_i$ the closure of $\operatorname{Lip}_b(X, \tau, \mathsf{d}_i)$ in $L^2(X, \mathfrak{m})$, we notice that $\mathcal{D}_i \subset \pi^i_*(L^2(X_i, \mathfrak{m}_i))$, because any function in $\operatorname{Lip}_b(X, \tau, \mathsf{d}_i)$ belongs to the image of π^i_*.

Denoting by $\widetilde{\mathrm{Ch}}_i$ and $\widetilde{\mathsf{P}}_t^i$ the Cheeger energy and its gradient flow in $(X_i, \tau_i, \widetilde{\mathsf{d}}_i, \mathfrak{m}_i)$, the formulas

$$\mathrm{Ch}_i(f) := \begin{cases} \widetilde{\mathrm{Ch}}_i(g) & \text{if } f = g \circ \pi^i \in \pi_*^i(L^2(X_i, \mathfrak{m}_i)) \\ +\infty & \text{otherwise} \end{cases} \tag{9.1}$$

$$\mathsf{P}_t^i f := (\widetilde{\mathsf{P}}_t^i g) \circ \pi^i \qquad \text{if } f = g \circ \pi^i \in \pi_*^i(L^2(X_i, \mathfrak{m}_i)), \quad t \geq 0 \,,$$

enable to lift the Cheeger energy $\widetilde{\mathrm{Ch}}_i$ and its gradient flow $\widetilde{\mathsf{P}}^i$ to the subspaces $\pi_*^i(L^2(X_i, \mathfrak{m}_i))$ of $L^2(X, \mathfrak{m})$, retaining the metric gradient flow property. Since $\widetilde{\mathrm{Ch}}_i$ have a dense domain in $L^2(X, \mathfrak{m}_i)$ it follows that the closure of the domain of Ch_i, namely $\pi_*^i(L^2(X_i, \mathfrak{m}_i))$, contains $\mathcal{D}_i$, so that Lemma 4.5 gives

$$\bigcup_{i \in I} D(\mathrm{Ch}_i) \text{ is dense in } L^2(X, \mathfrak{m}) \,. \tag{9.2}$$

The proof of the following theorem is inspired by various stability results based on Γ-convergence and on the energy dissipation point of view, see for instance [40], [24] and [25]. At the level of Ch_i, the only properties that will play a role are (9.2) and the energy dissipation inequality (9.4). The latter easily follows from the corresponding properties of $\widetilde{\mathrm{Ch}}_i$, $\widetilde{\mathsf{P}}^i$.

Theorem 9.1. *Under the previous assumptions on d_i one has that* Ch *coincides with the largest $L^2(X, \mathfrak{m})$ lower semicontinuous functional smaller than $\inf_i \mathrm{Ch}_i$.*

Proof. Let L_* be the largest $L^2(X, \mathfrak{m})$ lower semicontinuous functional smaller than $L := \inf_i \mathrm{Ch}_i$. Since $\mathsf{d}_i \leq \mathsf{d}$, from the inequality

$$\int |\mathrm{D}(g \circ \pi)|^2 \, \mathrm{d}\mathfrak{m} \leq \int |\mathrm{D}_{\mathsf{d}_i}(g \circ \pi)|^2 \, \mathrm{d}\mathfrak{m} = \int |\mathrm{D}_{\widetilde{\mathsf{d}}_i} g|^2 \, \mathrm{d}\mathfrak{m}_i \tag{9.3}$$

we immediately get $\mathrm{Ch} \leq \mathrm{Ch}_i$, hence $\mathrm{Ch} \leq L$ and the lower semicontinuity of Ch gives $\mathrm{Ch} \leq L_*$. In order to prove the converse inequality, we fix a probability density ρ with $C \geq \rho \geq c > 0$ $\mathfrak{m}$-a.e. in X and denote by ρ_t^i the gradient curves of Ch_i starting from ρ_0^i, the $L^2(X, \mathfrak{m})$ projection of ρ on $\overline{D(\mathrm{Ch}_i)}$. By (9.2) we know that $\rho_0^i \to \rho$ in $L^2(X, \mathfrak{m})$ and the stability of $\overline{D(\mathrm{Ch}_i)}$ under truncations immediately gives $C \geq \rho_t^i \geq c$ $\mathfrak{m}$-a.e. in X; in addition, using the regularization estimate (3.7) it is easily seen that

$$\limsup_{i \in I} \mathrm{Lip}(\rho_\cdot^i, (\delta, \infty)) < \infty \qquad \forall \delta > 0$$

(where the Lipschitz constant is computed w.r.t. the $L^2(X, \mathfrak{m})$ norm). Hence, we may find a subnet $\beta : J \to I$ and a curve ρ_t such that $\lim_{j \in J} \rho_t^{\beta(j)} = \rho_t$ in the weak topology of $L^2(X, \mathfrak{m})$ for all $t \geq 0$, with $\rho : (0, \infty) \to L^2(X, \mathfrak{m})$ continuous.

Our goal is to pass to the limit first w.r.t. j and then as $t \downarrow 0$ in the energy dissipation inequalities

$$\mathrm{Ent}(\mu_t^{\beta(j)}) + \frac{1}{2} \int_0^t |\dot{\mu}_s^{\beta(j)}|_{\beta(j)}^2 + \mathsf{F}^{\beta(j)}(\rho_s^{\beta(j)}) \, \mathrm{d}s \leq \mathrm{Ent}(\rho_0^{\beta(j)} \mathfrak{m}) \,, \tag{9.4}$$

with $\mu_t^{\beta(j)} = \rho_t^{\beta(j)} \mathfrak{m}$, $|\dot{\mu}_t^{\beta(j)}|_{\beta(j)}$ equal to the metric derivative of the curve $\mu_t^{\beta(j)}$ w.r.t. $W_{\mathsf{d}_{\beta(j)}}$ and $\mathsf{F}^{\beta(j)}$ the Fisher information functionals associated to $\mathrm{Ch}_{\beta(j)}$, to prove that ρ_t is the gradient curve of Ch starting from ρ.

We first notice that the representation (3.2) of the action as a supremum, together with the monotone convergence $\lim_j W_{\mathsf{d}_{\beta(j)}} = W_{\mathsf{d}}$ imply that $\mu_s \in AC^2([0, t]; (\mathscr{P}(X), W_{\mathsf{d}}))$ and that

$$\liminf_{j \in J} \int_0^t |\dot{\mu}_s^{\beta(j)}|_{\beta(j)}^2 \, \mathrm{d}s \geq \int_0^t |\dot{\mu}_s|^2 \, \mathrm{d}s \,, \tag{9.5}$$

where $\mu_t = \rho_t \mathfrak{m}$ and $|\dot{\mu}_t|$ denotes the metric derivative w.r.t. W_{d}.

Let us denote by F the Fisher information functional associated to Ch and notice that $\mathsf{F}_i \geq \mathsf{F}$. Hence, combining (9.5) with (9.4) and with $\liminf_{j \in J} \mathrm{Ent}(\mu_t^{\beta(j)}) \geq \mathrm{Ent}(\mu_t)$ we get

$$\mathrm{Ent}(\mu_t) + \frac{1}{2} \int_0^t |\dot{\mu}_s|^2 + \mathsf{F}(\rho_s) \, \mathrm{d}s \leq \mathrm{Ent}(\rho \mathfrak{m}) \,. \tag{9.6}$$

Since t is arbitrary this inequality, according to Theorem 8.2, proves that $\rho_t = \mathsf{P}_t\rho$, where P_t is the $L^2(X, \mathfrak{m})$-gradient flow of Ch; in addition, we can still use Theorem 8.2 to obtain that equality holds in (9.6). Therefore we obtain from this limiting argument the additional informations

$$\liminf_{j \in J} \text{Ent}(\rho_t^{\beta(j)}\mathfrak{m}) = \text{Ent}(\rho_t\mathfrak{m}) \,, \qquad \liminf_{j \in J} \int_0^t |\dot{\mu}_s^{\beta(j)}|_{\beta(j)}^2 \, \mathrm{d}s = \int_0^t |\dot{\mu}_s|^2 \, \mathrm{d}s \,, \qquad (9.7)$$

(that we shall exploit in the next theorem) as well as

$$\liminf_{j \in J} \int_0^t \mathsf{F}^{\beta(j)}(\rho_s^{\beta(j)}) \, \mathrm{d}s = \int_0^t \mathsf{F}(\rho_s) \, \mathrm{d}s \,. \qquad (9.8)$$

If we assume

$$\limsup_{t \downarrow 0} \frac{1}{t} \int_0^t \mathsf{F}(\rho_s) \, \mathrm{d}s \le \mathsf{F}(\rho) \qquad (9.9)$$

we can find, thanks to the convexity of $\mathsf{F}^{\beta(j)}$, $t(j) \to 0$ such that the functions

$$v_j := \frac{1}{t(j)} \int_0^{t(j)} \rho_s^{\beta(j)} \, \mathrm{d}s$$

satisfy $\liminf_j \mathsf{F}^{\beta(j)}(v_j) \le \mathsf{F}(\rho)$, so that

$$\liminf_{j \in J} \text{Ch}_{\beta(j)}(\sqrt{v_j}) \, \mathrm{d}s \le \text{Ch}(\sqrt{\rho}) \,. \qquad (9.10)$$

In order to prove that this implies $L_*(\sqrt{\rho}) \le \text{Ch}(\sqrt{\rho})$ it is sufficient to show that $\sqrt{v_j} \to \sqrt{\rho}$ in $L^2(X, \mathfrak{m})$. This can be proved as follows: since $W_{\mathsf{d}_{\beta(j)}}(v_j, \rho) \to 0$, we obtain

$$\lim_{j \in J} \int f v_j \, \mathrm{d}\mathfrak{m} = \int f \rho \, \mathrm{d}\mathfrak{m} \qquad \text{for all } f \in \bigcup_{i \in I} \text{Lip}(X, \tau, \mathsf{d}_i) \,.$$

Hence, by w^*-compactness of closed balls in $L^\infty(X, \mathfrak{m})$ and density of $\cup_i \text{Lip}(X, \tau, \mathsf{d}_i)$, $v_j \to \rho$ weakly* in $L^\infty(X, \mathfrak{m})$. Then, the entropy bound $\text{Ent}(v_j\mathfrak{m}) \le \text{Ent}(\rho\mathfrak{m})$ implies convergence of v_j in $\mathfrak{m}$-measure.

Now we remove the assumption (9.9). Given a probability density $\bar{\rho} \in D(\text{Ch})$ with $C \ge \bar{\rho} \ge c > 0$ $\mathfrak{m}$-a.e. in X we obtain by the previous step applied to $\rho = \mathsf{P}_t\bar{\rho}$, the inequality $\text{Ch}(\sqrt{\mathsf{P}_t\bar{\rho}}) \ge L_*(\sqrt{\mathsf{P}_t\bar{\rho}})$ for a.e. $t > 0$. By the chain rule, since $\text{Ch}(\mathsf{P}_t\bar{\rho}) \to \text{Ch}(\bar{\rho})$ as $t \downarrow 0$ implies $|\mathsf{D}\mathsf{P}_t\bar{\rho}|_w \to |\mathsf{D}\bar{\rho}|_w$ in $L^2(X, \mathfrak{m})$, we obtain $\text{Ch}(\sqrt{\mathsf{P}_t\bar{\rho}}) \to \text{Ch}(\sqrt{\bar{\rho}})$ as $t \downarrow 0$ and therefore $\text{Ch}(\sqrt{\bar{\rho}}) \ge L_*(\sqrt{\bar{\rho}})$.

This proves the inequality $\text{Ch} \ge L_*$ on all bounded and normalized functions ρ with $\inf \rho > 0$. Finally, we can extend by standard approximation arguments the inequality first to all bounded functions (by homogeneity and translation invariance) and eventually to all functions in $L^2(X, \mathfrak{m})$. $\qquad \square$

A byproduct of the previous proof and of the identification of gradient flows is the following stability result of gradient flows of Cheeger's energies; the stability proof provides also a crucial regularity property of Cheeger's energies that we call, as in [5], τ-upper regularity (see also Definition 12.4). We use the same notation of the statement of Theorem 9.1 and the notation

$$|\mathsf{D}(g \circ \pi^i)|_{w, \mathsf{d}_i} := |\mathsf{D}g|_w \circ \pi^i \qquad g \in D(\widetilde{\text{Ch}}_i) \,,$$

consistent with the definition (9.1) of Ch_j. It is not difficult to show, along the lines of (9.3), that $|\mathsf{D}f|_{w, \mathsf{d}_i} \ge |\mathsf{D}f|_{w, \mathsf{d}}$ $\mathfrak{m}$-a.e. in X for all $f \in D(\text{Ch}_i)$.

Theorem 9.2 (Stability of gradient flows and τ-upper regularity of Cheeger energies). *Under the same assumptions of Theorem 9.1, let $\rho_0 \in L_+^\infty(X, \mathfrak{m})$ and let ρ_t^i (resp. ρ_t) be the L^2 gradient curves of Ch_i (resp. Ch) starting from ρ_0^i, the $L^2(X, \mathfrak{m})$ projection of ρ_0 on $\overline{D(\text{Ch}_i)}$. Then $\rho_t^i \to \rho_t$ strongly in $L^2(X, \mathfrak{m})$ for all $t \ge 0$.*
In addition, for all $f \in D(\text{Ch})$ there exist a subnet $\beta : J \to I$, bounded and $\mathsf{d}_{\beta(j)}$-Lipschitz functions f_j with $f_j \to f$ in $L^2(X, \mathfrak{m})$ and $\text{Lip}_a(f_j, \mathsf{d}_{\beta(j)}, \cdot) \to |\mathsf{D}f|_w$ in $L^2(X, \mathfrak{m})$.

Proof. The weak convergence of ρ_t^i to ρ_t in $L^\infty(X, \mathfrak{m})$ has already been achieved in the proof of the previous theorem. To show that the convergence is actually strong, we use the first equality in (9.7), which can be improved to

$$\lim_{i \in I} \mathrm{Ent}(\rho_t^i \mathfrak{m}) = \mathrm{Ent}(\rho_t \mathfrak{m})$$

since it can be applied to an arbitrary subnet.

Also the last statement can be obtained with a small refinement of the proof of Theorem 9.1: it suffices to start from (9.10) and then to proceed as in the rest of the proof to obtain $k_j \in D(\mathrm{Ch}_{\beta(j)})$ with $k_j \to f$ in $L^2(X, \mathfrak{m})$ and $\liminf_j \mathrm{Ch}_{\beta(j)}(k_j) \leq \mathrm{Ch}(f)$. Now, the inequality $|\mathrm{D}k_j|_{w,\mathsf{d}_{\beta(j)}} \geq |\mathrm{D}k_j|_{w,\mathsf{d}}$ $\mathfrak{m}$-a.e. in X gives

$$\int \left(|\mathrm{D}k_j|_{w,\mathsf{d}_{\beta(j)}} - |\mathrm{D}f|_{w,\mathsf{d}} \right)^2 \mathrm{d}\mathfrak{m} \leq \mathrm{Ch}_{\beta(j)}(k_j) + \mathrm{Ch}(f) - 2 \int |\mathrm{D}k_j|_{w,\mathsf{d}}^2 \, \mathrm{d}\mathfrak{m} \ .$$

Hence, $|\mathrm{D}k_j|_{w,\mathsf{d}_{\beta(j)}} \to |\mathrm{D}f|_{w,\mathsf{d}}$ in $L^2(X, \mathfrak{m})$ along a further subnet. Finally, writing $k_j = \widetilde{k}_j \circ \pi_{\beta(j)}$, by applying Proposition 6.3(b) to $\widetilde{k}_j$ we can find $\widetilde{f}_j \in \mathrm{Lip}_b(X_{\beta(j)}, \widetilde{\mathsf{d}}_{\beta(j)})$ with

$$\lim_{j \in J} \int \left| |\mathrm{D}\widetilde{k}_j|_{w,\widetilde{\mathsf{d}}_{\beta(j)}} - \mathrm{Lip}_a(\widetilde{f}_j, \cdot) \right|^2 \mathrm{d}\mathfrak{m}_{\beta(j)} = 0 \ .$$

Setting $f_j = \widetilde{f}_j \circ \pi_{\beta(j)}$ we obtain the final part of the statement. $\qquad\square$

10. Energy measure spaces

In this section we study a class of extended distances in the framework of Dirichlet forms, basic references on this topic are [15], [23].

10.1. **Dirichlet forms, energy measure spaces and the Bakry-Émery condition.**

Definition 10.1 (Energy measure space). *We say that $(X, \mathcal{B}, \mathcal{E}, \mathfrak{m})$ is an energy measure space if:*

(a) *$\mathcal{B}$ is a σ-algebra in X and $\mathfrak{m} : \mathcal{B} \to [0,1]$ is a probability measure;*

(b) *$\mathcal{E}$ is a strongly local and Markovian Dirichlet form in $L^2(X, \mathfrak{m}) = L^2(X, \mathcal{B}, \mathfrak{m})$ whose domain*

$$\mathbb{V} = V := \{ f \in L^2(X, \mathfrak{m}) : \ \mathcal{E}(f, f) < \infty \}$$

is dense in $L^2(X, \mathfrak{m})$;

(c) *$\mathcal{E}$ admits a carré du champs operator defined on $\mathbb{V} \cap L^\infty(X, \mathfrak{m})$.*

Recall that the Markovian property means $\mathcal{E}(\phi \circ f, \phi \circ f) \leq \mathcal{E}(f, f)$ for all $f \in \mathbb{V}$ and all 1-Lipschitz $\phi : \mathbb{R} \to \mathbb{R}$. We recall that the carré du champs operator is the bilinear form $\Gamma : \left(\mathbb{V} \cap L^\infty(X, \mathfrak{m}) \right)^2 \to L^1(X, \mathfrak{m})$ providing a local representation of $\mathcal{E}$. When Γ exists (in more general situations it has to be understood as a measure-valued operator), it is characterized by the identity

$$\int \Gamma(f, f) \varphi \, \mathrm{d}\mathfrak{m} = -\frac{1}{2} \mathcal{E}(f^2, \varphi) + \mathcal{E}(f, f\varphi) \qquad \forall f, \varphi \in \mathbb{V} \cap L^\infty(X, \mathfrak{m}) \ . \tag{10.1}$$

We use the standard abbreviations $\mathcal{E}(f)$, $\Gamma(f)$ for $\mathcal{E}(f, f)$, $\Gamma(f, f)$, respectively, in the sequel. The domain $\mathbb{V}$ of $\mathcal{E}$ is endowed with the Hilbert norm

$$\|f\|_{\mathbb{V}}^2 := \|f\|_2^2 + \mathcal{E}(f) \tag{10.2}$$

and we denote by $\mathcal{A}_{\mathcal{E}}$ the Banach algebra $\{ f \in \mathbb{V} : \ f, \Gamma(f) \in L^\infty(X, \mathfrak{m}) \}$ endowed with the norm (see also (7.3))

$$\|f\|_{\mathcal{A}_{\mathcal{E}}} := \|f\|_\infty + \left\| \Gamma(f)^{1/2} \right\|_\infty \tag{10.3}$$

We now recall the main properties of the heat flow $\mathsf{P}^{\mathcal{E}}$ associated to $\mathcal{E}$. It can be defined as the unique locally absolutely continuous (in fact analytic) map $t \in (0, \infty) \mapsto f_t \in L^2(X, \mathfrak{m})$ satisfying

$$\frac{\mathrm{d}}{\mathrm{d}t} f_t = \Delta_{\mathcal{E}} f_t \quad \text{for a.e. } t \in (0, \infty) \ , \qquad \lim_{t \downarrow 0} f_t = f \quad \text{in } L^2(X, \mathfrak{m}) \ ,$$

where $\Delta_{\mathcal{E}} f$, the infinitesimal generator of the semigroup, is related to $\mathcal{E}$ by

$$v = \Delta_{\mathcal{E}} f \iff f \in \mathbb{V}, \ v \in L^2(X, \mathbf{m}), \quad \int vg \, \mathrm{d}\mathbf{m} = -\mathcal{E}(f, g) \quad \forall g \in \mathbb{V} \,.$$

Using the characterization

$$v = \Delta_{\mathcal{E}} f \iff -v \in \partial^- \frac{1}{2} \mathcal{E}(f)$$

it is easy to check that $\mathsf{P}_t^{\mathcal{E}}$ is also the metric gradient flow of $\frac{1}{2}\mathcal{E}$ with respect to the $L^2(X, \mathbf{m})$ distance, according to (EDI) and (EVI$_K$) with $K = 0$.

We recall now a few basic properties of $\mathsf{P}^{\mathcal{E}}$.

Since (thanks to the Markov property) $\mathsf{P}^{\mathcal{E}}$ is a contraction also in the $L^1(X, \mathbf{m})$ norm we can canonically extend it to a linear semigroup in $L^1(X, \mathbf{m})$, thanks to the density of $L^2(X, \mathbf{m})$ in $L^1(X, \mathbf{m})$. This extension of the semigroup, for which we retain the notation $\mathsf{P}^{\mathcal{E}}$, obviously satisfies

$$\int g \mathsf{P}_t^{\mathcal{E}} f \, \mathrm{d}\mathbf{m} = \int f \mathsf{P}_t^{\mathcal{E}} g \, \mathrm{d}\mathbf{m} \qquad \forall f \in L^1(X, \mathbf{m}), \ g \in L^\infty(X, \mathbf{m}), \ t \geq 0 \,. \tag{10.4}$$

Proposition 10.2 (Properties of $\mathsf{P}^{\mathcal{E}}$ and derivative of the entropy)**.** $\mathsf{P}^{\mathcal{E}}$ *is a Markov self-adjoint linear semigroup in* $L^2(X, \mathbf{m})$, $\Delta_{\mathcal{E}}$ *has a dense domain and* $t \mapsto \int \mathsf{P}_t^{\mathcal{E}} f \ln \mathsf{P}_t^{\mathcal{E}} f \, \mathrm{d}\mathbf{m}$ *is locally absolutely continuous in* $[0, \infty)$ *for all* $\mu = f \mathbf{m} \in D(\mathrm{Ent})$ *with*

$$-\frac{\mathrm{d}}{\mathrm{d}t} \int \mathsf{P}_t^{\mathcal{E}} f \ln \mathsf{P}_t^{\mathcal{E}} f \, \mathrm{d}\mathbf{m} = 4\mathcal{E}(\sqrt{\mathsf{P}_t^{\mathcal{E}} f}) = \int_{\{\mathsf{P}_t^{\mathcal{E}} f > 0\}} \frac{\Gamma(\mathsf{P}_t^{\mathcal{E}} f)}{\mathsf{P}_t^{\mathcal{E}} f} \, \mathrm{d}\mathbf{m} \qquad \text{for a.e. } t > 0 \,. \tag{10.5}$$

In addition, if $L^2(X, \mathbf{m})$ *is separable,* $\mathbb{V}$ *is a separable Hilbert space.*

Proof. The first properties are standard in the theory of semigroups, while (10.5) follows by the chain rule if $f \geq c > 0$ $\mathbf{m}$-a.e. in X and by an easy approximation, since $\mathbf{m}$ is finite, in the general case.

In order to prove separability, recall that, according to a standard results in the theory of semigroups (see for instance [4, Lem. 4.9]), it suffices to find a separable and $\mathsf{P}^{\mathcal{E}}$-invariant subspace $\mathbb{V}' \subset \mathbb{V}$. The subspace

$$\mathbb{V}' := \bigcup_{t > 0} \mathsf{P}_t^{\mathcal{E}} L^2(X, \mathbf{m})$$

is $\mathsf{P}^{\mathcal{E}}$-invariant. Its separability follows by the separability of $L^2(X, \mathbf{m})$ and from the regularizing estimate $\mathcal{E}(\mathsf{P}_t^{\mathcal{E}} f) \leq \|f\|_2^2 / t$ for all $t > 0$ and $f \in L^2(X, \mathbf{m})$, which corresponds to (3.7) with $K = 0$, $\bar{x} = f$ and $z = 0$. $\qquad\square$

We recall one of the possible formulation of the functional Bakry-Émery condition [11] for energy-measure spaces [5, 12]; other equivalent characterization in this abstract framework may be found in [5, Sect. 2.2], see also [10].

Definition 10.3 (Bakry-Émery condition via gradient contractivity)**.** *We say that the energy-measure space* $(X, \mathcal{B}, \mathcal{E}, \mathbf{m})$ *satisfies the Bakry-Émery condition* $\mathsf{BE}(K, \infty)$, $K \in \mathbb{R}$, *if*

$$\text{for every } g \in \mathcal{A}_{\mathcal{E}} \qquad \Gamma(\mathsf{P}_t^{\mathcal{E}} g) \leq \mathrm{e}^{-2Kt} \mathsf{P}_t^{\mathcal{E}} \Gamma(g) \quad \mathbf{m}\text{-a.e. in } X, \text{ for all } t \geq 0 \,. \tag{$\mathsf{BE}(K, \infty)$}$$

10.2. **Extended distances induced by an energy measure space.** In this context the definition of $\mathsf{CE}^2(X, \mathrm{Ch}, \mathbf{m})$ given in the metric setting can be immediately adapted, namely a curve ρ_s of probability densities belongs to $\mathsf{CE}^2(X, \mathcal{E}, \mathbf{m})$ if for some $c \in L^2(0, T)$ one has

$$\left| \int f \rho_t \, \mathrm{d}\mathbf{m} - \int f \rho_s \, \mathrm{d}\mathbf{m} \right| \leq \int_s^t c(r) \left(\int \Gamma(f) \rho_r \, \mathrm{d}\mathbf{m} \right)^{1/2} \mathrm{d}r \qquad \forall f \in \mathcal{A}_{\mathcal{E}} \tag{10.6}$$

for $0 \leq s \leq t \leq T$. The least c will still be denoted by $\|\rho_t'\|$.

Also the counterparts $W_{\mathcal{E}}$ and $W_{\mathcal{E}, *}$ of W_{Ch} and $W_{\mathrm{Ch}, *}$ can be immediately defined:

Definition 10.4. *For $\mu_0 = \rho_0\mathfrak{m}$, $\mu_1 = \rho_1\mathfrak{m} \in \mathscr{P}^a(X)$ we define*

$$W_{\mathcal{E}}^2(\mu_0, \mu_1) := \inf \left\{ \int_0^1 \|\rho_t'\|^2 \, \mathrm{d}t : \ \rho_t \in \mathsf{CE}^2(X, \mathcal{E}, \mathfrak{m}) \right\} \tag{10.7}$$

and

$$W_{\mathcal{E},*}^2(\mu_0, \mu_1) := 2 \sup_{\phi} \int (\phi_1\rho_1 - \phi_0\rho_0) \, \mathrm{d}\mathfrak{m} \,, \tag{10.8}$$

where the supremum runs among all $(\mathscr{L}^1 \otimes \mathcal{B})$-measurable bounded maps $\phi(t, x) = \phi_t(x)$ satisfying $\phi \in C_{\mathrm{w}*}([0, 1]; L^\infty(X, \mathfrak{m})) \cap L^1([0, 1]; \mathbb{V})$, and

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi_t + \frac{1}{2}\Gamma(\phi_t) \leq 0 \qquad \text{in } (0, 1) \times X, \text{ in the duality with } \mathcal{A}_{\mathcal{E}} \,. \tag{10.9}$$

As for the metric theory, we will use use often the simplified notation $W_{\mathcal{E}}(\rho_0, \rho_1)$, $W_{\mathcal{E},*}(\rho_0, \rho_1)$ for $W_{\mathcal{E}}(\rho_0\mathfrak{m}, \rho_1\mathfrak{m})$, $W_{\mathcal{E},*}(\rho_0\mathfrak{m}, \rho_1\mathfrak{m})$ respectively.

Arguing as in the proof of Proposition 7.4, it is easily seen that $W_{\mathcal{E}}$ is length and that $W_{\mathcal{E}}^2$ is jointly convex. In addition, with the same proof given in the metric setting, $t \mapsto \rho_t := \mathsf{P}_t\rho \in \mathsf{CE}^2(X, \mathcal{E}, \mathfrak{m})$ in any bounded interval $[0, T]$ with

$$\|\rho_t'\|^2 \leq \int_{\{\rho_t > 0\}} \frac{\Gamma(\rho_t)}{\rho_t} \, \mathrm{d}\mathfrak{m} \qquad \text{for a.e. } t > 0 \,. \tag{10.10}$$

Concerning $W_{\mathcal{E},*}$ one can also extend the same considerations of Lemma 7.7, obtaining in particular an equivalent definition where the supremum in (10.8) runs in $C^k([0, 1]; \mathcal{A}_{\mathcal{E}})$; as for Remark 7.8, it is also easy to check the joint convexity and the lower semicontinuity of $W_{\mathcal{E},*}^2$ with respect to the weak L^1-topology.

The following result can be obtained with the same proof given in the metric setting, see Proposition 7.9.

Proposition 10.5. $W_{\mathcal{E}} \geq W_{\mathcal{E},*}$ *on $\mathscr{P}^a(X) \times \mathscr{P}^a(X)$.*

Let us quickly discuss two cases when it is possible to prove that the distance $W_{\mathcal{E}}$ (and a fortiori $W_{\mathcal{E},*}$) between two probability densities is finite.

Lemma 10.6. *Let us suppose that $\mathcal{E}$ satisfies the global Poincaré inequality*

$$\int \left| f - \int f \, \mathrm{d}\mathfrak{m} \right|^2 \, \mathrm{d}\mathfrak{m} \leq \mathrm{c}_P \mathcal{E}(f) \quad \text{for every } f \in \mathbb{V} \tag{10.11}$$

Then if $\rho_0, \rho_1 \in L^2(X, \mathfrak{m})$ are probability densities with $\rho_i \geq \varrho > 0$ $\mathfrak{m}$-a.e. in X, $i = 0, 1$, we have

$$W_{\mathcal{E}}^2(\rho_0, \rho_1) \leq \frac{\mathrm{c}_P}{\varrho} \int |\rho_1 - \rho_0|^2 \, \mathrm{d}\mathfrak{m} \,. \tag{10.12}$$

Proof. We just take the linear connecting curve $\rho_s := (1 - s)\rho_0 + s\rho_1$, and we observe that, with $\bar{f} = \int f \, \mathrm{d}\mathfrak{m}$, for every $0 \leq s < t \leq 1$ one has

$$\frac{1}{t - s} \int f(\rho_t - \rho_s) \, \mathrm{d}\mathfrak{m} = \int (\rho_1 - \rho_0) f \, \mathrm{d}\mathfrak{m} = \int (\rho_1 - \rho_0)(f - \bar{f}) \, \mathrm{d}\mathfrak{m}$$

$$\leq \|\rho_1 - \rho_0\|_2 \|f - \bar{f}\|_2 \leq \left(\frac{\mathrm{c}_P}{\varrho}\right)^{1/2} \|\rho_1 - \rho_0\|_2 \left(\int \rho_r \Gamma(f) \, \mathrm{d}\mathfrak{m}\right)^{1/2} \,.$$

$\square$

Lemma 10.7. *Let us suppose that $\mathcal{E}$ satisfies the Logarithmic Sobolev inequality*

$$2 \operatorname{Ent}(\rho\mathfrak{m}) \leq \mathrm{c}_{LS} \int_{\{\rho > 0\}} \frac{\Gamma(\rho)}{\rho} \, \mathrm{d}\mathfrak{m} = 4\mathrm{c}_{LS}\mathcal{E}(\sqrt{\rho}) \tag{10.13}$$

for every probability density ρ with $\sqrt{\rho} \in \mathbb{V}$. Then the Talagrand inequality holds

$$\frac{1}{2}W_{\mathcal{E}}^2(\mu, \mathfrak{m}) \leq \mathrm{c}_{LS} \operatorname{Ent}(\mu) \quad \text{for every } \mu \in D(\operatorname{Ent}) \,. \tag{10.14}$$

In particular, $W_{\mathcal{E}}(\mu_0, \mu_1) < \infty$ whenever $\mu_i = \rho_i\mathfrak{m} \in D(\operatorname{Ent})$.

Proof. We follow the argument of [37] to use the Logarithmic-Sobolev inequality in order to show that the heat flow $\mathsf{P}_t^{\mathcal{E}}\mu$ connects μ to $\mathbf{m}$ and to estimate its length.

If $\mu = \rho\mathbf{m} \in D(\mathrm{Ent})$ the curve $\mu_t = \rho_t\mathbf{m}$ where $\rho_t = \mathsf{P}_t^{\mathcal{E}}\rho$ belongs to $\mathsf{CE}^2(X, \mathcal{E}, \mathbf{m})$ in any bounded interval $[0, T]$, since for every $f \in \mathcal{A}_{\mathcal{E}}$ and for every $0 \leq s \leq t$

$$\int f(\rho_t - \rho_s)\, \mathrm{d}\mathbf{m} = -\int_s^t \int \Gamma(f, \rho_r)\, \mathrm{d}\mathbf{m}\, \mathrm{d}r$$

$$\leq \int_s^t \Big(\int_{\{\rho_r>0\}} \frac{\Gamma(\rho_r)}{\rho_r}\, \mathrm{d}\mathbf{m}\Big)^{1/2} \Big(\int \rho_r \Gamma(f)\, \mathrm{d}\mathbf{m}\Big)^{1/2}\, \mathrm{d}r\ ; \tag{10.15}$$

the same formula shows that

$$\|\rho_t'\| \leq \Big(\int_{\{\rho_t>0\}} \frac{\Gamma(\rho_t)}{\rho_t}\, \mathrm{d}\mathbf{m}\Big)^{1/2}\ .$$

On the other hand, for every time t with $\mathrm{Ent}(\mu_t) > 0$, (10.5) yields

$$-\frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathrm{Ent}(\mu_t)\Big)^{1/2} = \int_{\rho_t>0} \frac{\Gamma(\rho_t)}{\rho_t}\, \mathrm{d}\mathbf{m} \cdot \Big(4\,\mathrm{Ent}(\mu_t)\Big)^{-1/2} \geq \Big(\frac{1}{2\mathrm{c}_{LS}} \int_{\{\rho_t>0\}} \frac{\Gamma(\rho_t)}{\rho_t}\, \mathrm{d}\mathbf{m}\Big)^{1/2}\ .$$
$$\tag{10.16}$$

Since the left hand side is integrable in $(0, \infty)$, also the right hand side is integrable and, in particular, the essential $\liminf$ of $\int_{\{\rho_t>0\}} \Gamma(\rho_t)/\rho_t\, \mathrm{d}\mathbf{m}$ as $t \to \infty$ is null. From (10.13) and the monotonicity of entropy we conclude that $\mathrm{Ent}(\mu_t) \to 0$ as $t \to \infty$ and thus $\rho_t \to 1$ in $L^1(X, \mathbf{m})$ as $t \to \infty$. Thus we can pass to the limit $t \to \infty$ in (10.15) and obtain that $t \mapsto \rho_t$ connects μ to $\mathbf{m}$ and is admissible (after reparametrization on a finite interval) in the definition of $W_{\mathcal{E}}$. Moreover, (10.16) provides the velocity estimate

$$\Big(\frac{1}{2\mathrm{c}_{LS}}\Big)^{1/2} \|\rho_t'\| \leq -\frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathrm{Ent}(\mu_t)\Big)^{1/2}\ .$$

Eventually, integrating this last inequality from 0 to ∞, recalling that $\mathrm{Ent}(\mu) \geq 0$ and that $\mathrm{Ent}(\mu_t) \to 0$ as $t \to \infty$ the Talagrand inequality (10.14) follows. $\qquad\square$

When $\mathsf{BE}(K, \infty)$ holds with $K > 0$, the well known argument of Bakry and Émery (we will also provide a proof based on the EVI formulation, which in turn follows by $\mathsf{BE}(K, \infty)$, see Corollary 11.5) yields the validity of the Logarithmic Sobolev inequality (10.13) with $\mathrm{c}_{LS} = K$, *provided* $(X, \mathcal{B}, \mathcal{E}, \mathbf{m})$ satisfies the *irreducibility condition*

$$f \in \mathbb{V},\ \mathcal{E}(f) = 0 \qquad \Longrightarrow \qquad f = c\ \mathbf{m}\text{-a.e. in } X \text{ for some } c \in \mathbb{R} \tag{10.17}$$

which is also equivalent to the L^2-*ergodicity* of the semigroup $\mathsf{P}^{\mathcal{E}}$:

$$\lim_{t\to\infty} \mathsf{P}_t^{\mathcal{E}} f = \int f\, \mathrm{d}\mathbf{m} \text{ strongly in } L^2(X, \mathbf{m}), \quad \text{for every } f \in L^2(X, \mathbf{m})\ . \tag{10.18}$$

Corollary 10.8. *If* $(X, \mathcal{B}, \mathcal{E}, \mathbf{m})$ *is irreducible according to* (10.17) *and* $\mathsf{BE}(K, \infty)$ *holds with* $K > 0$ *then* (10.13) *and* (10.14) *are satisfied with* $\mathrm{c}_{LS} = K$. *In particular every couple of probability measures* $\mu_i = \rho_i\mathbf{m} \in D(\mathrm{Ent})$ *has finite distance* $W_{\mathcal{E}}(\mu_0, \mu_1) < \infty$.

Our goal is now to prove that under the contractivity assumption $\mathsf{BE}(K, \infty)$ the upper length distance associated to $W_{\mathcal{E},*}$ coincides with $W_{\mathcal{E}}$. In the proof the following lemma will play a crucial role.

Lemma 10.9. *If* $\mathsf{BE}(K, \infty)$ *holds, then for any curve* $(\mu_t)_{t\in[0,1]}$ *in* $AC^2\big([0, 1], (\mathscr{P}(X), W_{\mathcal{E},*})\big)$ *with* $\mu_t = \rho_t\mathbf{m}$ *and any* $\phi \in \mathcal{A}_{\mathcal{E}}$ *one has (denoting by* $|\dot{\mu}_t|$ *the metric derivative w.r.t.* $W_{\mathcal{E},*}$*)*

$$\left|\frac{\mathrm{d}}{\mathrm{d}s}\int \rho_s\phi\, \mathrm{d}\mathbf{m}\right|_{s=t} \leq |\dot{\mu}_t| \cdot \sqrt{\int \Gamma(\phi)\rho_t\, \mathrm{d}\mathbf{m}} \qquad \text{for a.e. } t \in (0, 1)\ . \tag{10.19}$$

Proof. First, defining $W_{\mathcal{E},*,1} \leq W_{\mathcal{E},*}$ as in (7.15) of Remark 7.10, we obtain that $s \mapsto \int \phi \rho_s \, d\mathfrak{m}$ is absolutely continuous for all $\phi \in \mathcal{A}_{\mathcal{E}}$. Let $\psi \in \mathcal{A}_{\mathcal{E}} \cap D(\Delta_{\mathcal{E}})$ with $\inf \psi > 0$. Using the identity $-\psi^{-1}\Delta_{\mathcal{E}}\psi + \Delta_{\mathcal{E}} \log \psi = -\Gamma(\log \psi)$ and the gradient contractivity condition it is easy to check that, for $K \geq 0$,

$$\psi_s := 2\mathsf{P}_s^{\mathcal{E}} \log \mathsf{P}_{\delta-s}^{\mathcal{E}} \psi \qquad s \in [0,\delta]$$

is admissible in $W_{\mathcal{E},*}$. We fix a point t where $s \mapsto \mu_s$ is metrically differentiable and $s \mapsto \int \rho_s \log \psi \, d\mathfrak{m}$ is differentiable. Fix $s > t$ and $\delta = \lambda(s-t)$ with $\lambda > 0$. From the inequality

$$\delta\left|\int (\mathsf{P}_\delta^{\mathcal{E}} \log \psi)\rho_t \, d\mathfrak{m} - (\log \mathsf{P}_\delta^{\mathcal{E}} \psi)\rho_s \, d\mathfrak{m}\right| \leq \frac{1}{4}W_{\mathrm{Ch},*}^2(\mu_s, \mu_t) \, ,$$

dividing both sides by $(s-t)^2$ and letting $s \to t$ gives

$$\lambda\left|\lambda \int (\Delta_{\mathcal{E}} \log \psi - \frac{\Delta_{\mathcal{E}}\psi}{\psi})\rho_t \, d\mathfrak{m} - \frac{\mathrm{d}}{\mathrm{d}t}\int \rho_t \log \psi \, d\mathfrak{m}\right| \leq \frac{1}{4}|\dot{\mu}_t|^2 \, .$$

It follows that

$$|\frac{\mathrm{d}}{\mathrm{d}t}\int \rho_t \log \psi \, d\mathfrak{m}| \leq \lambda \int \Gamma(\log \psi)\rho_t \, d\mathfrak{m} + \frac{1}{4\lambda}|\dot{\mu}_t|^2 \, .$$

By a simple approximation, the same inequality holds for a.e. $t \in (0,1)$ for all $\psi \in \mathcal{A}_{\mathcal{E}}$ with $\inf \psi > 0$, i.e. removing the assumption $\psi \in D(\Delta_{\mathcal{E}})$. By minimizing w.r.t. λ and setting $\psi = \mathrm{e}^\phi$ we get the result.

In the case $K < 0$ we need to consider the reparameterization $2\mathsf{P}_{\theta(s)}^{\mathcal{E}} \log \mathsf{P}_{\theta(\delta)-\theta(s)}^{\mathcal{E}}\psi$, $s \in [0,\delta]$, where $\theta(0) = 0$ and $\theta'(s) = \mathrm{e}^{-2Ks}$. Since $\theta(s) = s + o(s)$ as $s \downarrow 0$, the same expansions above work with this modified function. $\qquad\square$

Remark 10.10. Notice that Lemma 10.9 and the next proposition could be reproduced even in the metric setting, since the proof of Lemma 10.9 used only the diffusion formula $\Delta_{\mathcal{E}}\phi(f) = \phi'(f)\Delta_{\mathcal{E}}f + \phi''(f)\Gamma(f)$, known to be true also in the metric setting (see [26, Prop. 4.11]), where Δ might be nonlinear. On the other hand, we preferred to state these results in this section because gradient contractivity conditions are expected to hold only in presence of quadratic energies. In the subclass of Minkowski spaces, it is known (see [36]) that contractivity of the heat flow w.r.t. to W_{d} holds if and only if the Minkowski structure is induced by an inner product.

Proposition 10.11 ($W_{\mathcal{E}}$ *is the upper length distance of* $W_{\mathcal{E},*}$)**.** *If* $\mathsf{BE}(K,\infty)$ *holds, then* $W_{\mathcal{E}}$ *is the upper length extended distance associated to* $W_{\mathcal{E},*}$ *according to* (3.9)*.*

Proof. Since $W_{\mathcal{E}}$ is length and larger than $W_{\mathcal{E},*}$, one inequality is obvious. Let us apply Lemma 10.9 to obtain

$$\left|\int f\rho_t \, d\mathfrak{m} - \int f\rho_s \, d\mathfrak{m}\right| \leq \int_s^t |\dot{\mu}_r|\left(\int \Gamma(f)\rho_r \, d\mathfrak{m}\right)^{1/2} \mathrm{d}r$$

for all $\mu_t = \rho_t\mathfrak{m} \in AC^2\big([0,1], (\mathscr{P}(X), W_{\mathcal{E},*})\big)$ and all $f \in \mathcal{A}_{\mathcal{E}}$. It follows that μ_t is absolutely continuous w.r.t. $W_{\mathcal{E}}$ and that

$$W_{\mathcal{E}}^2(\mu_1, \mu_0) \leq \int_0^1 |\dot{\mu}_t|^2 \, \mathrm{d}t \, ,$$

which provides, by the arbitrariness of μ_t, the converse inequality. $\qquad\square$

We conclude this section by proving that $\mathsf{BE}(K,\infty)$ allows for a further regularization in the definition of $W_{\mathcal{E},*}$. Let us first introduce the Banach space

$$D_\infty(\Delta_{\mathcal{E}}) := \Big\{f \in L^\infty(X,\mathfrak{m}) \cap D(\Delta_{\mathcal{E}}) : \ \Delta_{\mathcal{E}}f \in L^\infty(X,\mathfrak{m})\Big\} \qquad (10.20)$$

endowed with the graph norm $\|f\|_{D_\infty} := \|f\|_\infty + \|\Delta_{\mathcal{E}}f\|_\infty$ and let us recall (see [7]) that if $\mathsf{BE}(K,\infty)$ holds then $D_\infty(\Delta_{\mathcal{E}})$ is an algebra, continuously imbedded in $\mathcal{A}_{\mathcal{E}}$; in particular, there exists a constant $C_K > 0$ satisfying

$$\big\|\Gamma(f)\big\|_\infty \leq C_K \|f\|_\infty \|f\|_{D_\infty} \leq C_K \|f\|_{D_\infty}^2 \quad \text{for every } f \in D_\infty(\Delta_{\mathcal{E}}) \qquad (10.21)$$

Lemma 10.12. *The extended distance $W_{\mathcal{E},*}$ can be expressed by the duality formula (10.8) where the supremum runs among all $\phi \in C^\infty([0,1]; D_\infty(\Delta_{\mathcal{E}}))$.*

Proof. Arguing as in Lemma 7.7 it is not restrictive to take the supremum of (10.8) assuming $\phi \in C^\infty([0,1]; \mathcal{A}_{\mathcal{E}})$. We then set $\phi_t^\varepsilon := \mathsf{h}^\varepsilon \phi_t$ where h^ε is the mollification of $\mathsf{P}^{\mathcal{E}}$ introduced by (10.28). Since h^ε is a bounded linear operator $\mathcal{A}_{\mathcal{E}}$ to $D_\infty(\Delta_{\mathcal{E}})$, the curve $t \mapsto \phi_t^\varepsilon$ still belongs to $C^\infty([0,1]; D_\infty(\Delta_{\mathcal{E}}))$. On the other hand, the commutation property (10.29) shows that ϕ^ε is still a subsolution to (10.9). Since $\phi_t^\varepsilon \to \phi_t$ weakly$*$ in $L^\infty(X, \mathbf{m})$ as $\varepsilon \downarrow 0$, it is immediate to check that

$$\lim_{\varepsilon\downarrow 0} 2 \int \left(\rho_1 \phi_1^\varepsilon - \rho_0 \phi_0^\varepsilon\right) \, \mathrm{d}\mathbf{m} = 2 \int \left(\rho_1 \phi_1 - \rho_0 \phi_0\right) \, \mathrm{d}\mathbf{m} \, .$$

$\square$

10.3. Bakry-Émery condition and contractivity of the Heat semigroup. We say that $\mathsf{P}^{\mathcal{E}}$ is K-contractive w.r.t. $W_{\mathcal{E}}$ if

$$W_{\mathcal{E}}(\mathsf{P}_t^{\mathcal{E}} f, \mathsf{P}_t^{\mathcal{E}} g) \le \mathrm{e}^{-Kt} W_{\mathcal{E}}(f, g) \qquad \forall t \ge 0 \tag{10.22}$$

for all $f, g \in L^1_+(X, \mathbf{m})$ with $\int f \, \mathrm{d}\mathbf{m} = \int g \, \mathrm{d}\mathbf{m} = 1$.

The proof of the Bakry–Émery gradient estimate below will use some results of the theory developed recently in [9, Thm. 4.6] for the continuity equation in metric measure spaces. Its proof uses Hilbert space techniques and a vanishing viscosity argument, choosing as algebra of functions the set $\mathbb{V} \cap L^\infty(X, \mathbf{m})$.

Theorem 10.13. *Assume that $L^2(X, \mathbf{m})$ is separable. Let $V \in D(\Delta_{\mathcal{E}}) \cap L^\infty(X, \mathbf{m})$ with $\Delta_{\mathcal{E}} V \in L^\infty(X, \mathbf{m})$. Then for any $\bar{\rho} \in L^\infty(X, \mathbf{m})$ and any $T \in (0, \infty)$ there exists $\rho \in C^0_{w*}([0, T]; L^\infty(X, \mathbf{m}))$ with $\rho_0 = \bar{\rho}$ and*

$$\int f \rho_{s_1} \, \mathrm{d}\mathbf{m} - \int f \rho_{s_2} \, \mathrm{d}\mathbf{m} = \int_{s_1}^{s_2} \int \Gamma(f, V) \rho_r \, \mathrm{d}\mathbf{m} \, \mathrm{d}r \qquad \forall 0 \le s_1 \le s_2 \le T \tag{10.23}$$

for all $f \in \mathbb{V} \cap L^\infty(X, \mathbf{m})$. In particular $\rho \in \mathsf{CE}^2(X, \mathcal{E}, \mathbf{m})$ and the metric derivative of $s \mapsto \mu_s$ w.r.t. $W_{\mathcal{E}}$ can be estimated by

$$|\dot{\mu}_s|^2 \le \int \Gamma(V) \rho_s \, \mathrm{d}\mathbf{m} \qquad \text{for a.e. } s \in (0, T) \, . \tag{10.24}$$

Now we prove that the Bakry–Émery gradient estimate $\mathsf{BE}(K, \infty)$ is equivalent to K-contractivity of $\mathsf{P}^{\mathcal{E}}$ w.r.t. $W_{\mathcal{E}}$.

Theorem 10.14. *If $\mathsf{P}^{\mathcal{E}}$ is K-contractive w.r.t. to $W_{\mathcal{E}}$ and $L^2(X, \mathbf{m})$ is separable, then the Bakry-Emery condition $\mathsf{BE}(K, \infty)$ holds. Conversely, if $\mathsf{BE}(K, \infty)$ holds, then $\mathsf{P}^{\mathcal{E}}$ is K-contractive w.r.t. $W_{\mathcal{E}}$.*

Proof. Fix $t \ge 0$. We first consider $g \in \mathbb{V} \cap L^\infty(X, \mathbf{m})$ with $\Delta_{\mathcal{E}} g \in L^\infty(X, \mathbf{m})$. Fix a measure $\mu_0 = \rho_0 \mathbf{m} \in \mathscr{P}(X)$ with $\rho_0 \in L^\infty(X, \mathbf{m})$. Let us denote by $(\rho_s)_{s \in [0,1]}$ the solution to the continuity equation driven by the gradient of the constant (w.r.t. the time parameter s) potential $V := \mathsf{P}_t^{\mathcal{E}} g$ starting from ρ_0, given by Theorem 10.13. Now, let us first note that for any $h > 0$ one can apply (10.23) with $f = \mathsf{P}_t^{\mathcal{E}} g$ to get

$$\int \rho_h \mathsf{P}_t^{\mathcal{E}} g \, \mathrm{d}\mathbf{m} - \int \rho_0 \mathsf{P}_t^{\mathcal{E}} g \, \mathrm{d}\mathbf{m} = \int_0^h \int \Gamma(\mathsf{P}_t^{\mathcal{E}} g) \rho_s \, \mathrm{d}\mathbf{m} \, \mathrm{d}s \, . \tag{10.25}$$

On the other hand, putting $\mu_s^t = (\mathsf{P}_t^{\mathcal{E}} \rho_s)\mathbf{m}$ and denoting by $|\dot{\mu}_s^t|$ the metric derivative of the curve $s \mapsto \mu_s^t$, we can estimate:

$$\int \rho_h \mathsf{P}_t^{\mathcal{E}} g \, \mathrm{d}\mathbf{m} - \int \rho_0 \mathsf{P}_t^{\mathcal{E}} g \, \mathrm{d}\mathbf{m} = \int g \mathsf{P}_t^{\mathcal{E}} \rho_h \, \mathrm{d}\mathbf{m} - \int g \mathsf{P}_t^{\mathcal{E}} \rho_0 \, \mathrm{d}\mathbf{m} \leq \int_0^h |\dot{\mu}_s^t| \left(\int \Gamma(g) \mathsf{P}_t^{\mathcal{E}} \rho_s \, \mathrm{d}\mathbf{m} \right)^{\frac{1}{2}} \mathrm{d}s$$

$$\leq \ \mathrm{e}^{-Kt} \left(\int_0^h |\dot{\mu}_s|^2 \, \mathrm{d}s \right)^{\frac{1}{2}} \left(\int_0^h \int \Gamma(g) \mathsf{P}_t^{\mathcal{E}} \rho_s \, \mathrm{d}\mathbf{m} \, \mathrm{d}s \right)^{\frac{1}{2}}$$

$$\leq \ \mathrm{e}^{-Kt} \left(\int_0^h \int \Gamma(\mathsf{P}_t^{\mathcal{E}} g) \rho_s \, \mathrm{d}\mathbf{m} \, \mathrm{d}s \right)^{\frac{1}{2}} \left(\int_0^h \int \mathsf{P}_t^{\mathcal{E}} \Gamma(g) \rho_s \, \mathrm{d}\mathbf{m} \, \mathrm{d}s \right)^{\frac{1}{2}},$$

where we have used first the inequality $|\dot{\mu}_s^t| \leq \mathrm{e}^{-Kt} |\dot{\mu}_s|$ (derived by the Wasserstein contraction (10.22)) and then (10.24). Combining with (10.25) we obtain

$$\int_0^h \int \Gamma(\mathsf{P}_t^{\mathcal{E}} g) \rho_s \, \mathrm{d}\mathbf{m} \, \mathrm{d}s \leq \mathrm{e}^{-2Kt} \int_0^h \int \mathsf{P}_t^{\mathcal{E}} \Gamma(g) \rho_s \, \mathrm{d}\mathbf{m} \, \mathrm{d}s . \tag{10.26}$$

Dividing by h, letting $h \downarrow 0$ and using the weak* continuity of the curve ρ_s we finally get:

$$\int \Gamma(\mathsf{P}_t^{\mathcal{E}} g) \rho_0 \, \mathrm{d}\mathbf{m} \ \leq \ \mathrm{e}^{-2Kt} \int \mathsf{P}_t^{\mathcal{E}} \Gamma(g) \rho_0 \, \mathrm{d}\mathbf{m} . \tag{10.27}$$

By homogeneity, the same inequality holds for any $\rho_0 \in L_+^{\infty}(X, \mathbf{m})$. This clearly implies $\Gamma(\mathsf{P}_t^{\mathcal{E}} g) \leq \mathrm{e}^{-2Kt} \mathsf{P}_t^{\mathcal{E}} \Gamma(g)$ $\mathbf{m}$-a.e. in X.

To prove the assertion for arbitrary $g \in \mathbb{V} \cap L^{\infty}(X, \mathbf{m})$ we argue by approximation. Consider the following mollification of the semigroup, defined for $\varepsilon > 0$ and $f \in L^2(X, \mathbf{m})$ via:

$$\mathsf{h}^{\varepsilon} f = \int_0^{\infty} \frac{1}{\varepsilon} \eta \left(\frac{t}{\varepsilon} \right) \mathrm{e}^{(K \wedge 0) t} \mathsf{P}_t^{\mathcal{E}} f \, \mathrm{d}t , \tag{10.28}$$

with a non-negative kernel $\eta \in C_c^{\infty}(0, \infty)$ satisfying $\int_0^{\infty} \eta(t) \, \mathrm{d}t = 1$. It is easily seen that h^{ε} is a linear contraction in $\mathcal{A}_{\mathcal{E}}$ satisfying

$$\|\mathsf{h}^{\varepsilon} f\|_{\infty} \leq \|f\|_{\infty}, \quad \Gamma(\mathsf{h}^{\varepsilon} f) \leq \mathsf{h}^{\varepsilon} \Gamma(f), \quad \text{for every } f \in \mathcal{A}_{\mathcal{E}} \tag{10.29}$$

and

$$f \in L^{\infty}(X, \mathbf{m}) \quad \Rightarrow \quad \Delta_{\mathcal{E}} f \in L^{\infty}(X, \mathbf{m}), \quad \|\Delta_{\mathcal{E}} f\|_{\infty} \leq C_{\varepsilon} \|f\|_{\infty} \tag{10.30}$$

for some constant $C_{\varepsilon} > 0$. From the previous argument we thus obtain

$$\int \Gamma(\mathsf{P}_t^{\mathcal{E}} g_{\varepsilon}) \rho_0 \, \mathrm{d}\mathbf{m} \leq \mathrm{e}^{-2Kt} \int \mathsf{P}_t^{\mathcal{E}} \Gamma(g_{\varepsilon}) \rho_0 \, \mathrm{d}\mathbf{m} ,$$

for any $\rho_0 \in L_+^{\infty}(X, \mathbf{m})$. To conclude, it is sufficient to check that as $\varepsilon \downarrow 0$ we have $\mathsf{P}_t^{\mathcal{E}} g_{\varepsilon} \to \mathsf{P}_t^{\mathcal{E}} g$ and $g_{\varepsilon} \to g$ in $\mathbb{V}$ as $\varepsilon \to 0$. By convexity of $\mathcal{E}$ this in turn follows from the fact that for any $f \in \mathbb{V}$ we have $\mathcal{E}(\mathsf{P}_s^{\mathcal{E}} f - f) \to 0$ as $s \downarrow 0$.

Eventually we can prove the statement for any $g \in \mathbb{V}$ with a truncation argument.

In order to prove the converse statement, notice first that $\mathsf{BE}(K, \infty)$ implies that $\mathcal{A}_{\mathcal{E}}$ is invariant under the action of the semigroup. Then, recalling the definition (10.6) of $\mathsf{CE}^2(X, \mathcal{E}, \mathbf{m})$, (10.4) shows that $\mathsf{P}_t^{\mathcal{E}}$ maps curves $\rho_s \in \mathsf{CE}^2(X, \mathcal{E}, \mathbf{m})$ to curves $\sigma_s := \mathsf{P}_t^{\mathcal{E}} \rho_s \in \mathsf{CE}^2(X, \mathcal{E}, \mathbf{m})$ with $\|\sigma_s'\| \leq \mathrm{e}^{-Kt} \|\rho_s'\|$. By minimization we obtain the contractivity property.   □

We prove now, by standard methods, that $\mathsf{BE}(K, \infty)$ implies K-contractivity of $\mathsf{P}^{\mathcal{E}}$ also w.r.t. $W_{\mathcal{E}, *}$, a property that will also follow as a consequence of the EVI estimates of the next section. We don't know if the converse implication, known to be true for $W_{\mathcal{E}}$ under the separability assumption on $L^2(X, \mathbf{m})$, holds.

Proposition 10.15. *If* $\mathsf{BE}(K, \infty)$ *holds, then* $W_{\mathcal{E}, *}(\mathsf{P}_t^{\mathcal{E}} \rho_0, \mathsf{P}_t^{\mathcal{E}} \rho_1) \leq \mathrm{e}^{-Kt} W_{\mathcal{E}, *}(\rho_0, \rho_1)$ *for all* $t \geq 0$ *and* $\rho_0 \mathbf{m}, \rho_1 \mathbf{m} \in \mathscr{P}^a(X)$.

Proof. Take ϕ admissible in the definition of $W_{\mathcal{E},*}$ and note that $\psi = \mathrm{e}^{2Kt}\mathsf{P}_t^{\mathcal{E}}\phi$ is again admissible. Indeed, $\mathsf{BE}(K,\infty)$ gives

$$\frac{\mathrm{d}}{\mathrm{d}s}\psi + \frac{1}{2}\Gamma(\psi) \le \mathrm{e}^{2Kt}\mathsf{P}_t^{\mathcal{E}}\Big[\frac{\mathrm{d}}{\mathrm{d}s}\phi + \frac{1}{2}\Gamma(\phi)\Big] \le 0 \ .$$

Thus the definition of $W_{\mathcal{E},*}$ gives:

$$\frac{1}{2}W_{\mathcal{E},*}^2(\rho_0,\rho_1) \ge \int \psi_1\rho_1\,\mathrm{d}\mathfrak{m} - \int \psi_0\rho_0\,\mathrm{d}\mathfrak{m} = \mathrm{e}^{2Kt}\int\big(\phi_1\mathsf{P}_t^{\mathcal{E}}\rho_1\,\mathrm{d}\mathfrak{m} - \phi_0\mathsf{P}_t^{\mathcal{E}}\rho_0\big)\,\mathrm{d}\mathfrak{m}$$

and the statement follows by taking the supremum w.r.t. ϕ. $\qquad\qquad\square$

11. From gradient contractivity to EVI and consequences

In this section $(X,\mathcal{B},\mathcal{E},\mathfrak{m})$ is an energy measure space satisfying $\mathsf{BE}(K,\infty)$ for some $K \in \mathbb{R}$. The main result of this section is:

Theorem 11.1 ($\mathsf{P}^{\mathcal{E}}$ satisfies EVI_K relative to $W_{\mathcal{E},*}$)**.** *For all $\mu = \rho\mathfrak{m} \in \mathscr{P}^a(X)$, $\sigma \in D(\mathrm{Ent})$ with $W_{\mathcal{E},*}(\mu,\sigma) < \infty$ one has $\mathrm{Ent}(\mathsf{P}_t^{\mathcal{E}}\rho\,\mathfrak{m}) < \infty$, $W_{\mathcal{E},*}(\mathsf{P}_t^{\mathcal{E}}\rho\,\mathfrak{m},\sigma) < \infty$ for all $t > 0$ and (recall that $\frac{\mathrm{d}^+}{\mathrm{d}t}$ stands for upper right derivative)*

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\frac{1}{2}W_{\mathcal{E},*}^2(\mathsf{P}_t^{\mathcal{E}}\rho\,\mathfrak{m},\sigma) + \frac{K}{2}W_{\mathcal{E},*}^2(\mathsf{P}_t^{\mathcal{E}}\rho\,\mathfrak{m},\sigma) \le \big[\,\mathrm{Ent}(\sigma) - \mathrm{Ent}(\mathsf{P}_t^{\mathcal{E}}\rho\,\mathfrak{m})\big] \qquad \forall t \ge 0 \ . \qquad (11.1)$$

Before entering into the technical details, let us briefly explain the main idea of the proof.

First of all, thanks to the semigroup property of $\mathsf{P}_t^{\mathcal{E}}$, it is sufficient to prove an "integrated" version of (11.1), namely

$$\frac{1}{2}W_{\mathcal{E},*}^2(\mathsf{P}_t^{\mathcal{E}}\rho\,\mathfrak{m},\sigma) + t\,\mathrm{Ent}(\mathsf{P}_t^{\mathcal{E}}\rho\,\mathfrak{m}) \le \frac{t}{\mathrm{I}_{2K}(t)}\frac{1}{2}W_{\mathcal{E},*}^2(\rho\mathfrak{m},\sigma) + t\,\mathrm{Ent}(\sigma) \qquad \forall t > 0 \ . \qquad (11.2)$$

Indeed, the expansion $t/\mathrm{I}_{2K}(t) = 1 - Kt + o(t)$ and the lower semicontinuity of Ent provide (11.1) at $t = 0$, and the semigroup property provides the result for all positive times. Notice also that (11.2) implies all finiteness properties in the statement of the theorem.

We express the left-hand side by using a dual representation formula, obtained by combining (10.8) with the classical conjugate representation of the Entropy functional

$$\mathrm{Ent}(\rho\,\mathfrak{m}) = \sup_{\zeta \in L^{\infty}(X,\mathfrak{m})}\Big(\int \rho\zeta\,\mathrm{d}\mathfrak{m} - \int \mathrm{e}^{\zeta-1}\,\mathrm{d}\mathfrak{m}\Big) = \sup_{\zeta \in L^{\infty}(X,\mathfrak{m})\cap\mathbb{V}}\Big(\int \rho\zeta\,\mathrm{d}\mathfrak{m} - \int \mathrm{e}^{\zeta-1}\,\mathrm{d}\mathfrak{m}\Big) \ ,$$

where we have restricted the supremum to functions in $\mathbb{V}$ by a standard regularization argument (e.g. by applying (10.28)). After the simple transformation $\zeta = 1 + \psi/t$ yields

$$t\,\mathrm{Ent}(\rho\,\mathfrak{m}) - t = \sup_{\psi \in L^{\infty}(X,\mathfrak{m})\cap\mathbb{V}}\Big(\int \rho\psi\,\mathrm{d}\mathfrak{m} - t\int \mathrm{e}^{\psi/t}\,\mathrm{d}\mathfrak{m}\Big) \ . \qquad (11.3)$$

Replacing ρ with $\mathsf{P}_t^{\mathcal{E}}\rho$, adding the squared distance term, and using the symmetry of $\mathsf{P}^{\mathcal{E}}$, we end up with

$$\frac{1}{2}W_{\mathcal{E},*}^2(\mathsf{P}_t^{\mathcal{E}}\rho\,\mathfrak{m},\sigma) + t\,\mathrm{Ent}(\mathsf{P}_t^{\mathcal{E}}\rho\,\mathfrak{m}) - t = \sup_{(\phi_t),\psi}\Big(\int \rho\mathsf{P}_t^{\mathcal{E}}(\phi_1 + \psi)\,\mathrm{d}\mathfrak{m} - \int \phi_0\,\mathrm{d}\sigma - t\int \mathrm{e}^{\psi/t}\,\mathrm{d}\mathfrak{m}\Big) \ ,$$

$$(11.4)$$

where ϕ runs among subsolutions of (10.9) and ψ runs in $L^{\infty}(X,\mathfrak{m}) \cap \mathbb{V}$.

Let us now suppose that for every choice of $(\phi_s)_{s\in[0,1]}$ and $\psi \in L^{\infty}(X,\mathfrak{m}) \cap \mathbb{V}$ we can find a curve $(\psi_s)_{s\in[0,1]}$ in $\mathcal{A}_{\mathcal{E}}$ such that

$$\psi_1 = \psi \quad \text{and} \quad \frac{\mathrm{d}}{\mathrm{d}s}\mathsf{P}_{ts}^{\mathcal{E}}(\phi_s + \psi_s) + \frac{\mathrm{e}^{2Kts}}{2}\Gamma\big(\mathsf{P}_{ts}^{\mathcal{E}}(\phi_s + \psi_s)\big) \le 0 \ . \qquad (11.5)$$

Recalling (7.13) and (7.14) with $a = 0$, $b = 1$, $\vartheta(s) = e^{2Kts}$, such a curve provides the following upper bound for the term inside the "sup" in (11.4)

$$\int \rho \, \mathsf{P}_t^{\mathcal{E}}(\phi_1 + \psi) \, \mathrm{d}\mathbf{m} - \int \phi_0 \, \mathrm{d}\sigma - t \int e^{\psi/t} \, \mathrm{d}\mathbf{m}$$

$$= \int \rho \, \mathsf{P}_t^{\mathcal{E}}(\phi_1 + \psi_1) \, \mathrm{d}\mathbf{m} - \int (\phi_0 + \psi_0) \, \mathrm{d}\sigma + \int \psi_0 \, \mathrm{d}\sigma - t \int e^{\psi_0/t} \, \mathrm{d}\mathbf{m} + t \int \left(e^{\psi_0/t} - e^{\psi/t} \right) \mathrm{d}\mathbf{m}$$

$$\leq \frac{t}{\mathrm{I}_{2K}(t)} W_{\mathcal{E},*}^2(\rho \, \mathbf{m}, \sigma) + t \, \mathrm{Ent}(\sigma) - t + t \int \left(e^{\psi_0/t} - e^{\psi/t} \right) \mathrm{d}\mathbf{m} \,. \tag{11.6}$$

Expanding (11.5) and recalling that $e^{2Kts} \Gamma\left(\mathsf{P}_{ts}^{\mathcal{E}}(\phi_s + \psi_s) \right) \leq \mathsf{P}_{ts}^{\mathcal{E}} \Gamma(\phi_s + \psi_s)$ by the Bakry-Émery condition, we see that (11.5) is surely satisfied if

$$\mathsf{P}_{ts}^{\mathcal{E}} \left(\frac{\mathrm{d}}{\mathrm{d}s} \psi_s + t \Delta_{\mathcal{E}} \psi_s + \frac{1}{2} \Gamma(\psi_s) + t \Delta_{\mathcal{E}} \phi_s + \Gamma(\phi_s, \psi_s) \right) \leq 0 \tag{11.7}$$

where we used the fact that

$$\mathsf{P}_{ts}^{\mathcal{E}} \left(\frac{\mathrm{d}}{\mathrm{d}s} \phi_s + \frac{1}{2} \Gamma(\phi_s) \right) \leq 0$$

since ϕ is a subsolution of (10.9) and $\mathsf{P}_{ts}^{\mathcal{E}}$ is positivity preserving. This property and the non-negativity of $\Gamma(\psi_s)$ show that a candidate for (11.7) is provided by the backward Cauchy problem

$$\frac{\mathrm{d}}{\mathrm{d}s} \psi_s + t \Delta_{\mathcal{E}} \psi_s + \Gamma(\psi_s) + t \Delta_{\mathcal{E}} \phi_s + \Gamma(\phi_s, \psi_s) = 0, \quad s \in [0,1], \qquad \psi_1 = \psi \tag{11.8}$$

which can be reduced to the linear backward parabolic problem

$$\frac{\mathrm{d}}{\mathrm{d}s} \zeta_s + t \Delta_{\mathcal{E}} \zeta_s + \zeta_s \Delta_{\mathcal{E}} \phi_s + \Gamma(\phi_s, \zeta_s) = 0, \quad s \in [0,1], \qquad \zeta_1 := e^{\psi/t} \tag{11.9}$$

by applying the well known Hopf-Cole transformation

$$\zeta_s := e^{\psi_s/t} \,. \tag{11.10}$$

In conclusion, we have found that solving (11.9) and setting $\psi_s := t \log \zeta_s$ we get the bound (11.6). Miraculously enough, since

$$\int \zeta_s \Delta_{\mathcal{E}} \phi_s \, \mathrm{d}\mathbf{m} = - \int \Gamma(\phi_s, \zeta_s) \, \mathrm{d}\mathbf{m}, \quad t \int \Delta_{\mathcal{E}} \zeta_s \, \mathrm{d}\mathbf{m} = 0$$

equation (11.9) is mass preserving, so that

$$\int e^{\psi_0/t} \, \mathrm{d}\mathbf{m} = \int \zeta_0 \, \mathrm{d}\mathbf{m} = \int \zeta_1 \, \mathrm{d}\mathbf{m} = \int e^{\psi/t} \, \mathrm{d}\mathbf{m}$$

and with this particular choice the last integral term of (11.6) vanishes; since ϕ and ψ are arbitrary, we obtain (11.2).

Let us now check the technical details of the above argument. We divide the proof in a few steps: first of all, we will prove the existence of a sufficiently smooth solution to (11.9). We will then show that it takes values in a compact interval of $(0, \infty)$, so that it will not be difficult to check that the logarithmic transformation $\psi_s = t \log \zeta_s$ provides an admissible solution to (11.7).

Step 1: *for every $t > 0$, $\zeta_1 \in \mathbb{V}$ and $\phi \in C^1([0,1]; D_\infty(\Delta_{\mathcal{E}}))$, there exists a solution $\zeta \in W^{1,2}(0,1; L^2(X, \mathbf{m})) \cap L^2(0,1; D(\Delta_{\mathcal{E}}))$ (and thus in $C^0([0,1]; \mathbb{V})$) of (11.9).*

Reversing the time order setting $\widetilde{\zeta}_s := \zeta_{1-s}$, $\widetilde{\phi}_s := \phi_{1-s}$ and recalling the "integration by parts" formula

$$- \int (\Delta_{\mathcal{E}} \widetilde{\phi}_s) \, \widetilde{\zeta}_s \eta \, \mathrm{d}\mathbf{m} = \mathcal{E}(\widetilde{\phi}_s, \widetilde{\zeta}_s \eta) = \int \widetilde{\zeta}_s \Gamma(\widetilde{\phi}_s, \widetilde{\eta}) \, \mathrm{d}\mathbf{m} + \int \eta \Gamma(\widetilde{\phi}_s, \widetilde{\zeta}_s) \, \mathrm{d}\mathbf{m} \qquad \eta \in \mathbb{V} \tag{11.11}$$

which holds since $\Delta \phi_s, \Gamma(\phi_s) \in L^\infty(X, \mathbf{m})$, (11.9) is equivalent to the forward Cauchy problem

$$\frac{\mathrm{d}}{\mathrm{d}s} \widetilde{\zeta}_s - t \Delta_{\mathcal{E}} \widetilde{\zeta}_s - \widetilde{\zeta}_s \Delta_{\mathcal{E}} \widetilde{\phi}_s - \Gamma(\widetilde{\phi}_s, \widetilde{\zeta}_s) = 0, \quad s \in [0,1], \qquad \widetilde{\zeta}_0 := e^{\psi/t} \in L^2(X, \mathbf{m}) \tag{11.12}$$

which admits the variational formulation

$$\frac{\mathrm{d}}{\mathrm{d}s} \int \widetilde{\zeta}_s \eta \, \mathrm{d}\mathbf{m} + a_s(\widetilde{\zeta}_s, \eta) = 0 \quad \text{in } (0,1) \quad \text{for every } \eta \in \mathbb{V}, \tag{11.13}$$

where $(a_s)_{s\in[0,1]}$ is the continuous family of bounded bilinear forms in $\mathbb{V}\times\mathbb{V}$

$$a_s(\zeta,\eta) := t\mathcal{E}(\zeta,\eta) + \int \zeta\Gamma(\widetilde{\phi}_s,\eta)\,\mathrm{d}\mathfrak{m} \qquad \zeta,\eta\in\mathbb{V} \tag{11.14}$$

Since $C := \sup_s \|\Gamma(\phi_s)^{1/2}\|_\infty < \infty$, we get

$$\left| \int \zeta\Gamma(\widetilde{\phi}_s,\eta)\,\mathrm{d}\mathfrak{m} \right| \leq C\|\zeta\|_2\mathcal{E}(\eta)^{1/2} \tag{11.15}$$

and we easily prove that there exist $\lambda,\alpha > 0$ (depending on t) such that

$$a_s(\zeta,\zeta) + \lambda\|\zeta\|_2^2 \geq \alpha\|\zeta\|_{\mathbb{V}}^2 \tag{11.16}$$

A (unique) variational solution $\widetilde{\zeta} \in W^{1,2}(0,1;\mathbb{V}') \cap L^2(0,1;\mathbb{V})$ (and therefore continuous with values in $L^2(X,\mathfrak{m})$) then follows by applying J.L. Lions Theorem, see [33, Sect. 4.4, Thm. 4.1]. On the other hand, (11.12) and the uniform L^∞ bound on $\Delta_\mathcal{E}\phi_s$ show that

$$\frac{\mathrm{d}}{\mathrm{d}s}\widetilde{\zeta}_s - t\Delta_\mathcal{E}\widetilde{\zeta}_s \in L^2(0,1;L^2(X,\mathfrak{m}))\;;$$

since $-t\Delta_\mathcal{E}$ is the selfadjoint operator in $L^2(X,\mathfrak{m})$ associated to the symmetric Dirichlet form $t\mathcal{E}$ and since $\widetilde{\zeta}_0 \in \mathbb{V}$, the standard regularity results for variational evolution equation in Hilbert spaces (see, e.g. [16, Chap. III, Sect. 3]) yield $\zeta_s \in W^{1,2}(0,1;L^2(X,\mathfrak{m}))$. Eventually the equation (11.12) provides the $L^2(0,1;D(\Delta_\mathcal{E}))$ regularity.

Step 2. *Under the same assumptions of the previous step, if $|\Delta_\mathcal{E}\phi_s| \leq D$ $\mathfrak{m}$-a.e. for every $s \in [0,1]$ and $0 < \alpha \leq \zeta_1 \leq \beta < \infty$ $\mathfrak{m}$-a.e., then*

$$\alpha\mathrm{e}^{-D(1-s)} \leq \zeta_0 \leq \beta\mathrm{e}^{D(1-s)} \quad \text{$\mathfrak{m}$-a.e. in X for every } s\in[0,1] \tag{11.17}$$

We just observe that for every function $\theta \in C^1([0,1])$ the perturbed solution $\omega_s := \widetilde{\zeta}_s - \theta_s$ satisfies the equation

$$\frac{\mathrm{d}}{\mathrm{d}s}\omega_s - t\Delta_\mathcal{E}\omega_s - \omega_s\Delta_\mathcal{E}\widetilde{\phi}_s - \Gamma(\widetilde{\phi}_s,\omega_s) = f_s, \quad f_s = -(\theta'_s + \theta_s\Delta_\mathcal{E}\widetilde{\phi}_s), \quad s\in[0,1] \tag{11.18}$$

which can also be written as

$$\frac{\mathrm{d}}{\mathrm{d}s}\int \omega_s\eta\,\mathrm{d}\mathfrak{m} + a_s(\omega_s,\eta) = \int f_s\omega_s\,\mathrm{d}\mathfrak{m} \quad \text{for every } \eta\in\mathbb{V} \quad \text{a.e. in } (0,1). \tag{11.19}$$

Choosing $\theta_s = \beta\mathrm{e}^{Ds}$ we get $f_s \leq 0$ and $\omega_0 \leq 0$. Choosing $\eta_s := (\omega_s)_+$ in (11.19) and using the Leibniz rule (whose validity can easily be justified in this setting) we get

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}s}\int \eta_s^2\,\mathrm{d}\mathfrak{m} - \lambda\int \eta_s^2\,\mathrm{d}\mathfrak{m} \leq 0 \tag{11.20}$$

where we used the fact that

$$a_s(\omega,\omega_+) = a_s(\omega_+,\omega_+) \geq -\lambda\int (\omega_+)^2\,\mathrm{d}\mathfrak{m} \quad \text{for every } \omega\in\mathbb{V}$$

Since $\eta_0 = 0$, (11.20) yields $\eta_s = 0$ $\mathfrak{m}$-a.e. in X for every s; we thus obtain $\omega_s \leq 0$ and therefore $\widetilde{\zeta}_s \leq \beta\mathrm{e}^{Ds}$. The same argument, choosing $\theta_s := \alpha\mathrm{e}^{-Ds}$ and $\eta_s = (\omega_s)_-$ yields the other inequality $\widetilde{\zeta}_s \geq \alpha\mathrm{e}^{-Ds}$.

Step 3: *If $\phi \in C^1([0,1];D_\infty(\Delta_\mathcal{E}))$ is a subsolution to (10.9) then $\psi_s := t\log\zeta_s$ satisfy*

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathsf{P}_{ts}^\mathcal{E}(\phi_s + \psi_s) + \mathrm{e}^{2Kts}\frac{1}{2}\Gamma\big(\mathsf{P}_{ts}^\mathcal{E}(\phi_s + \psi_s)\big) \leq 0, \quad s\in[0,1] \tag{11.21}$$

The transformation $\psi_s := t\log\zeta_s$ is admissible thanks to the lower and upper bounds proved in the previous step; using the fact that

$$\Gamma(\phi_s,\psi_s) = \frac{t}{\zeta_s}\Gamma(\phi_s,\zeta_s), \qquad \Delta_\mathcal{E}\psi_s + t\Gamma(\psi_s) = \frac{t}{\zeta_s}\Delta_\mathcal{E}\zeta_s \quad \text{in } L^1(X,\mathfrak{m})$$

we obtain (11.8); notice that $\psi \in W^{1,2}(0,1;L^2(X,\mathfrak{m}))\cap C^0([0,1];\mathbb{V})$ and $\Delta_\mathcal{E}\psi \in L^1(0,1;L^1(X,\mathfrak{m}))$.

Since ϕ is a subsolution to (10.9) we get

$$\frac{\mathrm{d}}{\mathrm{d}s}(\phi_s + \psi_s) + t\Delta_{\mathcal{E}}(\phi_s + \psi_s) + \frac{1}{2}\Gamma(\phi_s + \psi_s) \leq 0, \quad s \in [0,1] \tag{11.22}$$

applying the positivity preserving $\mathsf{P}_{ts}^{\mathcal{E}}$ and observing that

$$\frac{\mathrm{d}}{\mathrm{d}s}\left(\mathsf{P}_{ts}^{\mathcal{E}}\varphi_s\right) = \mathsf{P}_{ts}^{\mathcal{E}}\frac{\mathrm{d}}{\mathrm{d}s}\varphi_s + t\Delta_{\mathcal{E}}\mathsf{P}_{ts}^{\mathcal{E}}\varphi_s = \mathsf{P}_{ts}^{\mathcal{E}}\left(\frac{\mathrm{d}}{\mathrm{d}s}\varphi_s + t\Delta_{\mathcal{E}}\varphi_s\right)$$

whenever $\varphi \in W^{1,2}(0,1; L^2(X,\mathfrak{m}))$ with $\Delta_{\mathcal{E}}\varphi \in L^1(0,1; L^1(X,\mathfrak{m}))$, we get

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathsf{P}_{ts}^{\mathcal{E}}(\phi_s + \psi_s) + \mathsf{P}_{ts}^{\mathcal{E}}\frac{1}{2}\Gamma(\phi_s + \psi_s) \leq 0, \quad s \in [0,1] \tag{11.23}$$

which yields (11.21) by the $\mathsf{BE}(K,\infty)$ gradient commutation property. $\qquad\square$

The following corollary is a direct consequence of (11.2), see also the metric regularization estimate (3.6).

Corollary 11.2 (LlogL regularization). *For any $\mu = \rho\mathfrak{m} \in \mathscr{P}^a(X)$ and $\sigma \in D(\mathrm{Ent})$ we have:*

$$\mathrm{Ent}(\mathsf{P}_t^{\mathcal{E}}\rho\,\mathfrak{m}) \leq \mathrm{Ent}(\sigma) + \frac{K}{\mathrm{e}^{2Kt}-1}W_{\mathcal{E},*}^2(\mu,\nu) \qquad \forall t > 0. \tag{11.24}$$

Also the following corollary is a direct consequence of Theorem 3.5 and of the fact that $W_{\mathcal{E}}$ is the upper length distance induced by $W_{\mathcal{E},*}$.

Corollary 11.3 ($\mathsf{P}^{\mathcal{E}}$ satisfies EVI_K relative to $W_{\mathcal{E}}$). *For all $\mu = \rho\mathfrak{m} \in \mathscr{P}^a(X)$, $\sigma \in D(\mathrm{Ent})$ with $W_{\mathcal{E}}(\mu,\sigma) < \infty$ one has $\mathrm{Ent}(\mathsf{P}_t^{\mathcal{E}}\rho\,\mathfrak{m}) < \infty$, $W_{\mathcal{E}}(\mathsf{P}_t^{\mathcal{E}}\rho\,\mathfrak{m},\sigma) < \infty$ for all $t > 0$ and*

$$\frac{\mathrm{d}^+}{\mathrm{d}t}\frac{1}{2}W_{\mathcal{E}}^2(\mathsf{P}_t^{\mathcal{E}}\rho\,\mathfrak{m},\sigma) + \frac{K}{2}W_{\mathcal{E}}^2(\mathsf{P}_t^{\mathcal{E}}\rho\,\mathfrak{m},\sigma) \leq \mathrm{Ent}(\sigma) - \mathrm{Ent}(\mathsf{P}_t^{\mathcal{E}}\rho\,\mathfrak{m}) \qquad \forall t \geq 0. \tag{11.25}$$

We can now obtain the geodesic property of $D(\mathrm{Ent})$ and the convexity of Ent, relative to $W_{\mathcal{E}}$. This provides a link with the theory developed independently by Lott-Villani and Sturm of synthetic lower bounds on the Ricci tensor, based on convexity properties of Ent (see [46]).

Theorem 11.4 (Geodesic convexity of the entropy functional). *$(D(\mathrm{Ent}), W_{\mathcal{E}}))$ is an extended geodesic metric space: for every couple of measures $\mu, \nu \in D(\mathrm{Ent})$ with $W_{\mathcal{E}}(\mu,\nu) < \infty$ there exists a $W_{\mathcal{E}}$-Lipschitz curve $\mu_t : [0,1] \to D(\mathrm{Ent})$ such that*

$$\mu_0 = \mu, \quad \mu_1 = \nu, \quad W_{\mathcal{E}}(\mu_s,\mu_t) = |t-s|W_{\mathcal{E}}(\mu,\nu), \quad s,t \in [0,1]. \tag{11.26}$$

In addition, the "finitary" length distance generated by $W_{\mathcal{E},*}$ according to (3.8) coincides with $W_{\mathcal{E}}$ on $D(\mathrm{Ent}) \times D(\mathrm{Ent})$ and Ent is K-convex on every curve as in (11.26):

$$\mathrm{Ent}(\mu_t) \leq (1-t)\mathrm{Ent}(\mu_0) + t\,\mathrm{Ent}(\mu_1) - \frac{K}{2}t(1-t)W_{\mathcal{E}}^2(\mu,\nu). \tag{11.27}$$

Finally, for all $\mu = \rho\,\mathfrak{m} \in D(\mathrm{Ent})$ the slope of the entropy coincides with the Fisher information

$$4\mathcal{E}(\sqrt{\rho}) = |\mathrm{D}_{W_{\mathcal{E}}}^- \mathrm{Ent}|^2(\rho\,\mathfrak{m}). \tag{11.28}$$

Proof. We are going to apply Corollary 3.6 with $X = D(\mathrm{Ent})$, $\mathsf{d} = W_{\mathcal{E},*}$, $F = \mathrm{Ent}$ and $\mathsf{S} = \mathsf{P}$ (we identify here measures with probability densities, as usual). We know from Theorem 11.1 that S provides a EVI_K-gradient flow of F in X, hence $(\mathscr{P}^a(X), W_{\mathcal{E},*,\ell})$ (where $W_{\mathcal{E},*,\ell}$ is the length distance associated to $W_{\mathcal{E},*}$ as in (3.8)) is a length space and the same holds for all sublevels $\{\mathrm{Ent} \leq c\}$, $c \in [0,\infty)$. On the other hand, since the sublevels are compact w.r.t. the weak $L^1(X,\mathfrak{m})$ topology (thanks to Dunford-Pettis theorem), we immediately obtain that $(\{\mathrm{Ent} \leq c\}, W_{\mathcal{E},*})$ are complete, thanks to the lower semicontinuity of $W_{\mathcal{E},*}$ w.r.t. the weak $L^1(X,\mathfrak{m})$ convergence. It follows by Corollary 3.6 that on $D(\mathrm{Ent}) \times D(\mathrm{Ent})$ the distance $W_{\mathcal{E},*,\ell}$ coincides with the upper length distance induced by $W_{\mathcal{E},*}$, namely $W_{\mathcal{E}}$ (by Proposition 10.11).

Thanks to compactness, the length properties of the sublevels can be improved to geodesic properties by the remarks made after Corollary 3.6. We can now use Corollary 11.3 to improve the EVI_K property from $W_{\mathcal{E},*}$ to $W_{\mathcal{E}}$, getting then the convexity of Ent along geodesics of $W_{\mathcal{E}}$.

We need only to prove the inequality $\geq$ in (11.28), since the converse inequality can be proved independently of curvature assumptions, see the proof of Lemma 8.1 in the metric setup and recall (10.10). We start from the observation that for any $\rho \in L^2_+(X, \mathbf{m})$ one has

$$4\mathcal{E}(\sqrt{\mathsf{P}^{\mathcal{E}}_t \rho}) = |\mathrm{D}^-_{W_{\mathcal{E}}} \mathrm{Ent}\,|^2(\mathsf{P}^{\mathcal{E}}_t \rho\, \mathbf{m}) \qquad \text{for a.e. } t > 0$$

by looking at the energy dissipation rates from the L^2 point of view and by the $W_{\mathcal{E}}$ point of view (the latter is derived from the EVI_K property). Now, if $\rho \in \mathbb{V}$ with $\inf \rho > 0$ we can pass to the limit as $t \downarrow 0$ along a suitable sequence and use the lower semicontinuity of $|\mathrm{D}^-_{W_{\mathcal{E}}} \mathrm{Ent}\,|$ w.r.t. $W_{\mathcal{E}}$ convergence (derived from the convexity of entropy) to obtain the inequality $\geq$ in (11.28). For general probability densities we argue by truncation, using once more the lower semicontinuity of $|\mathrm{D}^-_{W_{\mathcal{E}}} \mathrm{Ent}\,|$. $\qquad\square$

We conclude pointing out some standard consequences of the K-convexity of Ent.

Corollary 11.5 (Convexity of bounded densities, Log-Sobolev and transport inequalities).
*(i) If $K \geq 0$, then the sets $\{\mu = \rho\, \mathbf{m} \in \mathscr{P}^a(X) : \|\rho\|_\infty \leq c\}$ are geodesically convex w.r.t. $W_{\mathcal{E}}$, i.e. every couple of measures $\mu_i = \rho_i \mathbf{m}$, $i = 0, 1$, with $\|\rho_i\|_\infty \leq c$ and $W_{\mathcal{E}}(\mu_0, \mu_1) < \infty$, can be connected by a geodesic $\mu_t = \rho_t \mathbf{m}$ as in (11.26) such that $\|\rho_t\|_\infty \leq c$ for every $t \in [0, 1]$.
(ii) If $K > 0$ and $\mathcal{E}$ is irreducible according to (10.17), then the log-Sobolev inequality (10.13) holds with $c_{LS} = K$. In particular, thanks to Lemma 10.7, one has the Talagrand inequality*

$$\frac{K}{2} W^2_{\mathcal{E}}(\mu, \mathbf{m}) \leq \mathrm{Ent}(\mu) \qquad \forall \mu \in \mathscr{P}^a(X)\,. \tag{11.29}$$

Proof. For the geodesic convexity of the sets $\{\mu = \rho\, \mathbf{m} \in \mathscr{P}^a(X) : \|\rho\|_\infty \leq c\}$ the rescaling of transport plans as in [4, Prop. 3.3] applies (notice that the assumptions on the side of the supports made therein play a role only when $K < 0$).

The metric argument of [2, Lem. 2.4.13], relying on K-convexity, can be applied to give

$$\mathrm{Ent}(\rho_1\, \mathbf{m}) - \mathrm{Ent}(\rho_0\, \mathbf{m}) \leq \frac{1}{2K} |\mathrm{D}^- \mathrm{Ent}\,|^2(\rho_1\, \mathbf{m})\,,$$

for any pair of measures $\mu_0 = \rho_0\, \mathbf{m}$, $\mu_1 = \rho_1\, \mathbf{m}$ in $D(\mathrm{Ent})$ with $W_{\mathcal{E}}(\mu_0, \mu_1) < \infty$. It $\rho \in D(\mathrm{Ent})$ we apply this inequality with $\rho_0 = \mathsf{P}_t \rho$ and $\rho_1 = \rho$ and we let $t \to \infty$ to get, by the irreducibility of $\mathcal{E}$ and (10.18),

$$\mathrm{Ent}(\rho\, \mathbf{m}) \leq \frac{1}{2K} |\mathrm{D}^- \mathrm{Ent}\,|^2(\rho\, \mathbf{m})\,.$$

Using (11.28) we obtain the log-Sobolev inequality. $\qquad\square$

Corollary 11.6. *The transport inequality (11.29) implies that the class of measures in $\mathscr{P}^a(X)$ with bounded density w.r.t. $\mathbf{m}$ is dense in $D(\mathrm{Ent})$ w.r.t. $W_{\mathcal{E}}$.*

Proof. Let us approximate μ by the measures $\mu^k := \mathbf{1}_{\{\rho \leq k\}} \mu / a_k$, where $a_k = \int_{\{\rho \leq k\}} \rho\, \mathrm{d}\mathbf{m} \uparrow 1$ are the normalization constants. If ρ^k are the densities of μ^k, writing $b_k = 1 - a_k = \int_{\{\rho > k\}} \rho\, \mathrm{d}\mathbf{m}$ and

$$\rho^k = \mathbf{1}_{\{\rho \leq k\}} \rho + \left(\frac{1}{a_k} - 1\right) \mathbf{1}_{\{\rho \leq k\}} \rho = \mathbf{1}_{\{\rho \leq k\}} \rho + b_k \rho^k\,, \qquad \rho = \mathbf{1}_{\{\rho \leq k\}} \rho + \mathbf{1}_{\{\rho > k\}} \rho$$

adding the constant term $\mathbf{1}_{\{\rho \leq k\}} \rho$ to the solutions to the continuity inequality we get

$$W^2_{\mathcal{E}}(\mu^k, \mu) \leq b_k W^2_{\mathcal{E}}\left(\mu^k, \frac{1}{b_k} \mathbf{1}_{\{\rho > k\}} \rho \mathbf{m}\right)\,.$$

It suffices then to show that $\lim_k b_k W_{\mathcal{E}}(\mu^k, b_k^{-1} \mathbf{1}_{\{\rho > k\}} \rho \mathbf{m}) = 0$. To this aim, we compare both measures with $\mathbf{m}$. The transport inequality then gives

$$b_k W^2_{\mathcal{E}}(\mu^k, \mathbf{m}) \leq \frac{2b_k}{K} \mathrm{Ent}(\mu^k) \to 0$$

and

$$b_k W^2_{\mathcal{E}}\left(\frac{1}{b_k} \mathbf{1}_{\{\rho > k\}} \rho, \mathbf{m}\right) \leq \frac{2}{K} \left[\int_{\{\rho > k\}} \rho \ln \rho\, \mathrm{d}\mathbf{m} + b_k \ln\left(\frac{1}{b_k}\right)\right] \to 0\,. \tag{11.30}$$

$\qquad\square$

12. FROM DIFFERENTIABLE TO METRIC STRUCTURES AND CONVERSELY

12.1. **Energy measure spaces induce extended metric measure spaces.** In this section $(X, \mathcal{B}, \mathcal{E}, \mathfrak{m})$ is an energy measure space according to Definition 10.1; following the construction explained in Section 4, page 13, we are going to introduce an extended metric-topological structure starting from given a family $\mathcal{L}$ of *pointwise defined* real functions such that

$$\mathcal{L} \subset \{f : X \to \mathbb{R} : f \text{ is } \mathcal{B}\text{-measurable and bounded}, \Gamma(f) \leq 1\},$$
$$\mathcal{L} \text{ separates points of } X, \tag{12.1a}$$

so that (equivalence $\mathfrak{m}$-a.e. classes of) elements of $\mathcal{L}$ belong to $\mathcal{A}_{\mathcal{E}}$. Then

$$\tau \text{ is the Hausdorff topology in } X \text{ generated by } \mathcal{L}, \tag{12.1b}$$

i.e. τ is the coarsest topology such that all the functions of $\mathcal{L}$ are continuous: (X, τ) is automatically completely regular. Restricting $\mathfrak{m}$ to $\mathscr{B}(\tau) \subset \mathcal{B}$, we will assume that

$$\mathfrak{m} \in \mathscr{P}(X) \quad (\text{i.e. } \mathfrak{m} \text{ is Radon in } \mathscr{B}(\tau)), \quad \operatorname{supp} \mathfrak{m} = X; \tag{12.1c}$$

we can then consider the class

$$\mathcal{A}_{\mathcal{E}}^* = \{f \in \mathcal{A}_{\mathcal{E}} \cap C_b(X, \tau) : \Gamma(f) \leq 1\} \tag{12.1d}$$

(where we identify functions in $\mathcal{A}_{\mathcal{E}}^*$ with their unique τ-continuous representative) containing $\mathcal{L}$ and use $\mathcal{A}_{\mathcal{E}}^*$ to define canonically $\mathsf{d}_{\mathcal{E}} : X \times X \to [0, \infty]$ by

$$\mathsf{d}_{\mathcal{E}}(x, y) := \sup\{|f(x) - f(y)| : f \in \mathcal{A}_{\mathcal{E}}^*\}, \tag{12.1e}$$

so that $(X, \mathsf{d}_{\mathcal{E}})$ is an extended metric space.

In addition, if I is the collection of finite subsets of $\mathcal{A}_{\mathcal{E}}^*$, for $i \in I$ we define

$$\mathsf{d}_i(x, y) := \sup_{f \in i} |f(x) - f(y)|. \tag{12.2}$$

Notice that d_i is only a semidistance, i.e. it is symmetric and it satisfies the triangle inequality. We shall also use the fact that $\mathsf{d}_i(\cdot, y) \in \mathcal{A}_{\mathcal{E}}^*$ for all $y \in X$, with $\Gamma(\mathsf{d}_i(\cdot, y)) \leq 1$ $\mathfrak{m}$-a.e. in X. If we endow X with the semidistance d_i it is immediately seen that (X, d_i) is separable.

We thus get that $(X, \tau, \mathsf{d}_{\mathcal{E}})$ is an extended metric-topological space according to Definition 4.1 and that $(X, \tau, \mathsf{d}_{\mathcal{E}}, \mathfrak{m})$ is an extended metric measure space according to Definition 4.7.

Remark 12.1. The typical case of this construction occurs when $\operatorname{supp} \mathfrak{m} = X$ and $\mathcal{L}$ can be identified with a subset of $\mathcal{A}_{\mathcal{E}} \cap C_b(X, \tau_0)$ for some preexisting topology τ_0 in X; in this case the condition $\operatorname{supp} \mathfrak{m} = X$ provides uniqueness of the continuous representative and τ is coarser than τ_0, so that (12.1c) is satisfied if $\mathfrak{m} \in \mathscr{P}(X, \tau_0)$. Notice that $\tau = \tau_0$ if τ_0 is generated by $\mathcal{L}$.

Proposition 12.2. *Under assumptions* (12.1a), (12.1c) *one has* $W_{\mathcal{E},*} \geq W_{\mathsf{d}_{\mathcal{E}}}$, *where* $W_{\mathcal{E},*}$ *is defined in* (10.8).

Proof. Let d_i as above. By Theorem 5.1 we need only to show that $W_{\mathsf{d}_i} \leq W_{\mathcal{E},*}$. In order to prove this property, taking (5.3) and the comments immediately after into account, it suffices to show that

$$Q_t^i \phi(x) := \inf_{y \in X} \phi(y) + \frac{1}{2t} \mathsf{d}_i^2(x, y)$$

is admissible in (7.10) whenever ϕ is bounded and d_i-Lipschitz. By applying Lemma 12.3 below to $Q_t^i \phi(\cdot)$ we get $\Gamma(Q_t^i \phi(\cdot)) \leq |\mathrm{D}_i Q_t^i \phi(\cdot)|^2$ $\mathfrak{m}$-a.e. in X, where $|\mathrm{D}_i f|$ denotes the slope w.r.t. d_i. Taking into account the subsolution property (7.7) of $Q_t^i \phi$, we obtain $\partial_t Q_t^i \phi + \Gamma(Q_t^i \phi)/2 \leq 0$. $\square$

Let d be a finite semidistance in X. In the proof of the next lemma we are going to use in (X, d) the following links between the descending slope in (3.4) computed w.r.t. d and the functions $Q_t f$, $f \in \operatorname{Lip}(X, \mathsf{d})$, provided by the Hopf-Lax formula

$$Q_t f(x) = \inf_{y \in Y} f(y) + \mathsf{d}^2(x, y)/2t, \tag{12.3}$$

see [5, Sec. 3] for the proof (see also [2, Lem. 3.1.5]):

$$|\mathrm{D}f|^2(x) \geq |\mathrm{D}^- f|^2(x) = \limsup_{t\downarrow 0} \int_0^1 \left(\frac{\mathrm{D}^+ f(x,tr)}{tr}\right)^2 \mathrm{d}r \ . \tag{12.4}$$

Here, $\mathrm{D}^+ f(x,t)$ is defined by

$$\mathrm{D}^+ f(x,t) := \sup \left\{ \limsup_{n\to\infty} \mathsf{d}(x_n, x) : \ (x_n) \text{ minimizing sequence in (12.3)} \right\} \ .$$

It is not hard to show, by diagonal arguments, that $\mathrm{D}^+ f$ is upper semicontinuous in $X \times (0,\infty)$, endowed with the product of d and of the Euclidean distance (see again [5, Sec. 3]). These results are stated in [5] for metric spaces, and they can be immediately adapted to degenerate space (X, d), just noticing that $\mathsf{d}(x,y) = 0$ implies $f(x) = f(y)$, i.e. lifting them from the quotient metric space to (X, d).

Lemma 12.3. *Let d be a bounded $(\tau \times \tau)$-continuous semidistance in X with (X, d) separable, $\mathsf{d}(\cdot, y) \in \mathbb{V}$ and $\Gamma(\mathsf{d}(\cdot, y)) \leq 1$ $\mathfrak{m}$-a.e. in X for all $y \in X$. Then, for all $f : X \to \mathbb{R}$ bounded and d-Lipschitz, denoting by $|\mathrm{D}f|$ the slope w.r.t. d, one has $\Gamma(f) \leq |\mathrm{D}f|^2$ $\mathfrak{m}$-a.e. in X.*

Proof. In the proof, which follows closely [5, Lem. 3.12], we will use the following weak stability property of the Γ operator, which follows easily by Mazur's lemma: if $f_n \in \mathbb{V}$ and $f_n \to f$ in $L^2(X, \mathfrak{m})$, then

$$\sqrt{\Gamma(f_n)} \to G \quad \text{weakly in } L^2(X, \mathfrak{m}) \qquad \text{implies} \qquad G \geq \sqrt{\Gamma(f)} \quad \mathfrak{m}\text{-a.e. in } X \ . \tag{12.5}$$

If (z_i) is a countable d-dense subset of X we define

$$Q_t^n f(x) = \min_{1 \leq i \leq n} f(z_i) + \frac{1}{2t}\mathsf{d}^2(z_i, x) \ , \qquad Q_t f(x) = \min_{y \in X} f(y) + \frac{1}{2t}\mathsf{d}^2(y, x) \ , \tag{12.6}$$

and we set $I_n(x) := \left\{ i \in \{1, \ldots, n\} : \ z_i \text{ minimizes (12.6)} \right\}$. By the density of (z_i), it is clear that $Q_t^n f \downarrow Q_t f$ as $n \to \infty$. Therefore, if $\zeta_n(x) \in I_n(x)$, it turns out that $(\zeta_n(x))$ is a minimizing sequence for $Q_t f(x)$, namely

$$\frac{1}{2t}\mathsf{d}^2(x, \zeta_n(x)) + f(\zeta_n(x)) \to Q_t f(x) \quad \text{as } n \to \infty \ .$$

The very definition of $\mathrm{D}^+ f(x,t)$ then gives

$$\limsup_{n\to\infty} \mathsf{d}(x, \zeta_n(x)) \leq \mathrm{D}^+ f(x,t) \ . \tag{12.7}$$

Since $Q_t^n f(x) = f(z_i) + \mathsf{d}^2(z_i, x)/2t$ on $\{x : \ \zeta_n(x) = z_i\}$, the locality property and the fact that $\mathsf{d}(z_i, \cdot)$ belongs to $\mathcal{A}_{\mathcal{CE}}$ together with the chain rule yield

$$\Gamma(Q_t^n f)(x) \leq \frac{1}{t^2} \max_{i \in I_n(x)} \mathsf{d}^2(x, z_i) \quad \text{for } \mathfrak{m}\text{-a.e. } x \in \{\zeta_n = z_i\} \ .$$

If we define $\zeta_n(x)$ as the smallest index j, among those that realize the maximum for $\mathsf{d}(z_i, x)$, $i \in I_n(x)$, the previous formula yields

$$\Gamma(Q_t^n f)(x) \leq \frac{1}{t^2}\mathsf{d}^2(x, \zeta_n(x)) \qquad \text{for } \mathfrak{m}\text{-a.e. } x \in X \ . \tag{12.8}$$

Since $Q_t^n f$ and $\Gamma(Q_t^n f)$ are uniformly bounded and $Q_t^n f$ converges pointwise to $Q_t f$, considering any weak limit point G of $\sqrt{\Gamma(Q_t^n f)}$ in $L^2(X, \mathfrak{m})$ we obtain by (12.5), (12.7) and (12.8) that

$$\Gamma(Q_t f)(x) \leq G^2(x) \leq \frac{\left(\mathrm{D}^+ f(x,t)\right)^2}{t^2} \quad \text{for } \mathfrak{m}\text{-a.e. } x \in X \ . \tag{12.9}$$

Since f is Lipschitz, it follows that $\mathrm{D}^+ f(x,t)/t$ is uniformly bounded and since d is $(\tau \times \tau)$-continuous the function $\mathrm{D}^+ f$ is Borel in $X \times (0,\infty)$. Integrating (12.4) on an arbitrary Borel

set A and applying Fatou's Lemma, from (12.9) we get

$$
\begin{aligned}
\int_A |\mathrm{D}f|^2 \, \mathrm{d}\mathfrak{m} &\geq \int_A \limsup_{t\downarrow 0} \int_0^1 \Big(\frac{D^+ f(x,tr)}{tr}\Big)^2 \, \mathrm{d}r \, \mathrm{d}\mathfrak{m}(x) \\
&\geq \limsup_{t\downarrow 0} \int_0^1 \int_A \Big(\frac{D^+ f(x,tr)}{tr}\Big)^2 \, \mathrm{d}\mathfrak{m}(x) \, \mathrm{d}r \\
&\geq \limsup_{t\downarrow 0} \int_0^1 \int_A \Gamma(Q_{tr}f)(x) \, \mathrm{d}\mathfrak{m}(x) \, \mathrm{d}r \\
&\geq \int_0^1 \liminf_{t\downarrow 0} \Big(\int_A \Gamma(Q_{tr}f) \, \mathrm{d}\mathfrak{m}\Big) \, \mathrm{d}r \geq \int_A \Gamma(f) \, \mathrm{d}\mathfrak{m} \,,
\end{aligned}
$$

where in the last inequality we applied (12.5) once more. Since A is arbitrary we conclude. $\quad\square$

The τ-upper regularity has already been identified in [5] as a crucial compatibility condition between the topological and the metric/differentiable structures, needed to identify $\mathcal{E}$ with a Cheeger energy.

Definition 12.4 (τ-upper regularity). *Let $(X,\mathcal{B},\mathcal{E},\mathfrak{m})$ be an energy measure space with $\mathscr{B}(\tau) \subset \mathcal{B}$ for some topology τ in X. We say that $\mathcal{E}$ is τ-upper regular if for all $f \in \mathbb{V}$ there exist:*

(a) *functions $f_n \in \mathrm{Lip}_b(X,\mathsf{d},\tau)$ with $f_n \to f$ in $L^2(X,\mathfrak{m})$;*

(b) *bounded τ-upper semicontinuous functions g_n with $g_n \geq \sqrt{\Gamma(f_n)}$ $\mathfrak{m}$-a.e. in X with $\limsup_n \int g_n^2 \, \mathrm{d}\mathfrak{m} \leq \mathcal{E}(f)$.*

Theorem 12.5. *Let $(X,\mathcal{B},\mathcal{E},\mathfrak{m})$ be an energy measure space and let $(X,\tau,\mathsf{d}_\mathcal{E},\mathfrak{m})$ be the extended metric-topological structure associated to a set $\mathcal{L}$ as in (12.1a,b,c,d,e). Then $\mathcal{E} \leq \mathrm{Ch}_{\mathsf{d}_\mathcal{E}}$ and, in particular, $W_\mathcal{E} \geq W_{\mathrm{Ch}_{\mathsf{d}_\mathcal{E}}}$. The equality $\mathcal{E} = \mathrm{Ch}_{\mathsf{d}_\mathcal{E}}$ holds iff $\mathcal{E}$ is τ-upper regular. In particular, if $\mathcal{E}$ is τ-upper regular, one has:*

(a) *the classes of 2-absolutely continuous curves $\mu_t = \rho_t \mathfrak{m}$ w.r.t. $W_\mathcal{E}$, $W_{\mathcal{E},*}$ and $W_{\mathsf{d}_\mathcal{E}}$ with $\rho_t \in L^\infty(L^\infty(X,\mathfrak{m}))$ coincide and the same is true for the corresponding metric derivatives;*

(b) *If $\mathsf{BE}(K,\infty)$ holds, the metric gradient flows of Ent w.r.t. $W_\mathcal{E}$, $W_{\mathcal{E},*}$ and $W_{\mathsf{d}_\mathcal{E}}$ coincide with $\mathsf{P}_t^\mathcal{E}$;*

(c) *If $\mathsf{BE}(K,\infty)$ holds and $\mathsf{P}_t^\mathcal{E}$ maps $C_b(X)$ in $C_b(X)$, then $W_\mathcal{E}$ is the upper length distance in $\mathscr{P}^a(X)$ associated to $W_{\mathsf{d}_\mathcal{E}}$ according to (3.9).*

(d) *If $\mathsf{BE}(K,\infty)$ holds with $K \geq 0$ (resp. $K > 0$ and P is irreducible), then $W_\mathcal{E}$ restricted to $\{\mu = \rho\,\mathfrak{m} : \|\rho\|_\infty \leq c\}$ (resp. $D(\mathrm{Ent})$) is the upper length distance in $\mathscr{P}^a(X)$ associated to $W_{\mathsf{d}_\mathcal{E}}$ according to (3.9).*

Proof. We prove first the inequality $\mathcal{E} \leq \mathrm{Ch}_{\mathsf{d}_\mathcal{E}}$. Taking Theorem 9.1 and the lower semicontinuity of $\mathcal{E}$ into account, it suffices to show that $\mathcal{E} \leq \mathrm{Ch}_i$, where Ch_i is the Cheeger energy associated to the semimetric measure space $(X,\tau,\mathsf{d}_i,\mathfrak{m})$ and d_i is the monotone approximation of $\mathsf{d}_\mathcal{E}$ illustrated by (12.2). By Lemma 12.3 we obtain

$$
\mathcal{E}(f) \leq \int |\mathrm{D}_i f|^2 \, \mathrm{d}\mathfrak{m}
$$

for all $f : X \to \mathbb{R}$ bounded and d_i-Lipschitz, where $|\mathrm{D}_i f|$ denotes the slope w.r.t. d_i. Using Proposition 6.3(b) and once more the lower semicontinuity of $\mathcal{E}$ we conclude.

The necessity of τ-upper regularity for the validity of the equality $\mathcal{E} = \mathrm{Ch}_{\mathsf{d}_\mathcal{E}}$ follows by applying Theorem 9.2 to the $(\tau\times\tau)$-continuous semidistances d_i which monotonically converge to $\mathsf{d}_\mathcal{E}$: one obtains the τ-upper regularity (along a subnet $i = \beta(j)$) with $f_i \in \mathrm{Lip}(X,\tau,\mathsf{d}_i)$ and with the d_i-upper semicontinuous (and thus also τ-upper semicontinuous) functions $g_i = \mathrm{Lip}_a(f_i,\mathsf{d}_i,\cdot)$.

For the sufficiency of τ-upper regularity we follow the argument in [5, Prop. 3.11]. Thanks to the τ-upper regularity, in order to prove that $\mathrm{Ch}_{\mathsf{d}_\mathcal{E}} \leq \mathcal{E}$ it is sufficient to show that $f \in D(\mathrm{Ch}_{\mathsf{d}_\mathcal{E}})$ and that $|\mathrm{D}f|_{w,\mathsf{d}_\mathcal{E}} \leq g$ $\mathfrak{m}$-a.e. in X whenever $f \in \mathrm{Lip}(X,\tau,\mathsf{d}_\mathcal{E})$ and g is a bounded τ-upper semicontinuous function such that $g \geq \sqrt{\Gamma(f)}$ $\mathfrak{m}$-a.e. in X. By the very definition of $\mathrm{Ch}_{\mathsf{d}_\mathcal{E}}$, $f \in D(\mathrm{Ch}_{\mathsf{d}_\mathcal{E}})$. In order to prove the inequality $|\mathrm{D}f|_{w,\mathsf{d}_\mathcal{E}} \leq g$ $\mathfrak{m}$-a.e. in X we will prove the

inequality for the slope $|\mathrm{D}f|_{\mathsf{d}_{\mathcal{E}}}$. We need only, thanks to the upper semicontinuity of g, to prove the pointwise inequality $|\mathrm{D}f|_{\mathsf{d}_{\mathcal{E}}} \leq c$ in the τ-open set $U := \{g < c\}$. By homogeneity, we can assume $c = 1$ and we fix $x_0 \in U$; since τ is generated by $\mathcal{L}$ we can find a finite collection $(f_n)_{n=1}^N$ of elements of $\mathcal{L}$ and $r > 0$ such that

$$F := \Big\{ x \in X : \max_{1 \leq n \leq N} |f_n(x) - f_n(x_0)| \leq r \Big\} \subset U \ .$$

Set

$$\delta(x) := \max_{1 \leq n \leq N} |f_n(x) - f_n(x_0)|, \quad l(x) := \min\{r, |f(x) - f(x_0)|\} \ , \quad h(x) := \max\{l(x), \delta(x)\} \ ,$$

and notice that $\delta \in \mathcal{A}_{\mathcal{E}}^*$. Since $\{h = l\} = \{l \geq \delta\} \subset F \subset U$ and since $\Gamma(l) \leq 1$ in U, by locality we get $\Gamma(h) \leq 1$ $\mathfrak{m}$-a.e. in X; since h is τ-continuous we get $h \in \mathcal{A}_{\mathcal{E}}^*$, so that (12.1e) yields

$$h(x) = h(x) - h(x_0) \leq \mathsf{d}_{\mathcal{E}}(x, x_0) \ .$$

Since the topology induced by $\mathsf{d}_{\mathcal{E}}$ is stronger than τ, the τ-continuity of f gives $h(x) \geq |f(x) - f(x_0)|$ for $\mathsf{d}_{\mathcal{E}}(x, x_0)$ sufficiently small. It follows that $|\mathrm{D}f|_{\mathsf{d}_{\mathcal{E}}}(x_0) \leq 1$.

Finally, statements (a) and (b) follow by Corollary 7.5 and Theorem 8.3 of the metric theory, taking also the inequalities $W_{\mathsf{d}_{\mathcal{E}}} \leq W_{\mathcal{E},*} \leq W_{\mathcal{E}}$ into account.

Let us prove now statement (c). It suffices to show that $W_{\mathcal{E}}^2(\rho_0 \mathfrak{m}, \rho_1 \mathfrak{m}) \leq \int_0^1 |\dot{\mu}_s|^2 \, \mathrm{d}s$ for any absolutely continuous curve μ_t w.r.t. $W_{\mathsf{d}_{\mathcal{E}}}$ contained in $\mathscr{P}^a(X)$, where $|\dot{\mu}_t|$ denotes the metric derivative w.r.t. $W_{\mathsf{d}_{\mathcal{E}}}$. Since $\mathcal{E}$ is τ-upper regular, we can identify $\mathcal{E}$ with $\mathrm{Ch}_{\mathsf{d}_{\mathcal{E}}}$ and $\Gamma(f)$ with $|\mathrm{D}f|_w^2$. By the definition of $W_{\mathcal{E}}$, it will be sufficient to prove the inequality

$$\left| \int \phi \rho_{s_1} \, \mathrm{d}\mathfrak{m} - \int \phi \rho_{s_2} \, \mathrm{d}\mathfrak{m} \right| \leq \int_{s_1}^{s_2} \left(\int \Gamma(\phi) \rho_s \, \mathrm{d}\mathfrak{m} \right)^{1/2} |\dot{\mu}_s| \, \mathrm{d}s \tag{12.10}$$

for all $\phi \in \mathcal{A}_{\mathcal{E}}$ and $0 \leq s_1 \leq s_2 \leq 1$.

We start from the observation that for any $\psi \in \mathcal{A}_{\mathcal{E}} \cap C_b(X)$ and any bounded τ-upper semicontinuous function $g \geq \sqrt{\Gamma(\psi)}$ $\mathfrak{m}$-a.e. in X the Lipschitz property of ψ w.r.t. $\mathsf{d}_{\mathcal{E}}$ can be "localized" as in the first part of the proof of the theorem to get that $\psi \circ \eta$ is absolutely continuous in $[0, 1]$ for all $\eta \in AC([0, 1]; (X, \mathsf{d}_{\mathcal{E}}))$ with $|(\psi \circ \eta)'| \leq g(\eta)|\dot{\eta}|$ a.e. in $(0, 1)$. By integrating this inequality along a test plan $\boldsymbol{\eta}$ representing the curve μ_t we get

$$\left| \int \psi \rho_{s_1} \, \mathrm{d}\mathfrak{m} - \int \psi \rho_{s_2} \, \mathrm{d}\mathfrak{m} \right| \leq \int_{s_1}^{s_2} \left(\int g^2 \rho_s \, \mathrm{d}\mathfrak{m} \right)^{1/2} |\dot{\mu}_s| \, \mathrm{d}s \ .$$

Now, as in [5, Thm. 3.15] we can apply the regularization property

$$\mathsf{P}_t^{\mathcal{E}} : L^\infty(X, \mathfrak{m}) \to \mathcal{A}_{\mathcal{E}} \ , \qquad \|\Gamma(\mathsf{P}_t^{\mathcal{E}} f)\|_\infty \leq c(K, t, \|f\|_\infty) \quad \forall f \in L^\infty(X, \mathfrak{m})$$

derived from $\mathsf{BE}(K, \infty)$ with Γ-calculus techniques (see for instance [5, Cor. 2.3] for a proof) and the Feller property $\mathsf{P}_t^{\mathcal{E}} : C_b(X) \to C_b(X)$ (which implies, by monotone approximation, that the class of bounded τ-upper semicontinuous functions is invariant under $\mathsf{P}_t^{\mathcal{E}}$) to get from the previous inequality with $\psi = \mathsf{P}_t^{\mathcal{E}} \phi$ the inequality

$$\left| \int \mathsf{P}_t^{\mathcal{E}} \phi \rho_{s_1} \, \mathrm{d}\mathfrak{m} - \int \mathsf{P}_t^{\mathcal{E}} \phi \rho_{s_2} \, \mathrm{d}\mathfrak{m} \right| \leq \int_{s_1}^{s_2} \left(\int (c \wedge \mathrm{e}^{-2Kt} \mathsf{P}_t^{\mathcal{E}} g^2) \rho_s \, \mathrm{d}\mathfrak{m} \right)^{1/2} |\dot{\mu}_s| \, \mathrm{d}s$$

with $c = c(K, t, \|\phi\|_\infty)$. Now, for all $\phi \in \mathcal{A}_{\mathcal{E}}$ we exploit once more the τ-upper regularity assumption, using in the previous inequality functions $\phi_n \in \mathcal{A}_{\mathcal{E}}$ with $\|\phi_n\|_\infty \leq \|\phi\|_\infty$ and bounded τ-upper semicontinuous functions $g_n \geq \sqrt{\Gamma(\phi_n)}$ with $\phi_n \to \phi$ and $g_n \to \sqrt{\Gamma(\phi)}$ in $L^2(X, \mathfrak{m})$, to get

$$\begin{aligned}
\left| \int \mathsf{P}_t^{\mathcal{E}} \phi \rho_{s_1} \, \mathrm{d}\mathfrak{m} - \int \mathsf{P}_t^{\mathcal{E}} \phi \rho_{s_2} \, \mathrm{d}\mathfrak{m} \right| &\leq \int_{s_1}^{s_2} \left(\int (c \wedge \mathrm{e}^{-2Kt} \mathsf{P}_t^{\mathcal{E}} \Gamma(\phi)) \rho_s \, \mathrm{d}\mathfrak{m} \right)^{1/2} |\dot{\mu}_s| \, \mathrm{d}s \\
&\leq \mathrm{e}^{-Kt} \int_{s_1}^{s_2} \left(\int \mathsf{P}_t^{\mathcal{E}} \Gamma(\phi) \rho_s \, \mathrm{d}\mathfrak{m} \right)^{1/2} |\dot{\mu}_s| \, \mathrm{d}s \ .
\end{aligned}$$

Eventually we can take the limit as $t \downarrow 0$ to obtain (12.10).

The proof of (d) in the case $K \geq 0$ is similar and uses the convexity properties of $\{\mu = \rho\,\mathfrak{m} : \|\rho\|_\infty \leq c\}$ to avoid the regularization based on the Feller property $\mathsf{P}_t^{\mathcal{E}} : C_b(X) \to C_b(X)$. In the case $K > 0$ the result can be extended to $D(\mathrm{Ent})$ thanks to the $W_{\mathcal{E}}$-density of $\{\mu = \rho\,\mathfrak{m} : \rho \in L^\infty(X, \mathfrak{m})\}$ in $D(\mathrm{Ent})$, ensured by Corollary 11.6. $\qquad\square$

12.2. **Extended metric measure spaces induce energy measure spaces.** In view of the results of this section, it is useful to consider the case when Ch is a quadratic form, namely to assume that the parallelogram identity holds:

$$\mathrm{Ch}(f + g) + \mathrm{Ch}(f - g) = 2\,\mathrm{Ch}(f) + 2\,\mathrm{Ch}(g) \text{ for all } f, g \in L^2(X, \mathfrak{m}) . \tag{12.11}$$

Definition 12.6 (Asymptotically Hilbertian spaces)**.** *We say that an extended metric measure space $(X, \tau, \mathsf{d}, \mathfrak{m})$ is asymptotically Hilbertian if Ch satisfies the parallelogram identity (12.11).*

In the proof of the next theorem we will also need the following calculus property, borrowed from [4].

Lemma 12.7 (Plan representing the gradient, horizontal and vertical derivatives)**.** *Assume that (X, d) is complete and let $u \in D(\mathrm{Ch}) \cap L^\infty(X, \mathfrak{m})$. Then there exists a test plan $\boldsymbol{\sigma} \in \mathscr{P}(X^{[0,1]})$ representing the gradient of u in the following sense:*

$$\lim_{t\downarrow 0} \frac{E_t}{t} = \lim_{t\downarrow 0} \frac{u \circ \mathrm{e}_0 - u \circ \mathrm{e}_t}{E_t} = |\mathrm{D}u|_w \circ \mathrm{e}_0 \quad \text{in } L^2(X^{[0,1]}, \boldsymbol{\sigma}) ,$$

where $E_t(\eta) := \sqrt{t \int_0^t |\dot\eta(s)|^2 \, \mathrm{d}s}$. Moreover, for any $f \in D(\mathrm{Ch})$ one has:

$$\liminf_{t\downarrow 0} \int \frac{f(\eta(t)) - f(\eta(0))}{t} \, \mathrm{d}\boldsymbol{\sigma} \geq \limsup_{\varepsilon\downarrow 0} \int \frac{|\mathrm{D}u|_w^2(\eta(0)) - |\mathrm{D}(u + \varepsilon f)|_w^2(\eta(0))}{2\varepsilon} \, \mathrm{d}\boldsymbol{\sigma} . \tag{12.12}$$

Proof. The first result is proven as in [4, Lem. 4.15] and relies on Proposition 5.5 applied to the semigroup P_t starting from $v := ce^u$ (with c suitable normalization constant), and then defining $\boldsymbol{\sigma} := v^{-1} \circ e_0\boldsymbol{\eta}$. The possibility to apply Proposition 5.5 to the gradient flow is ensured by (6.4) and Theorem 7.3. The second result is obtained as in [4, Lem. 4.5]. $\qquad\square$

Theorem 12.8. *If $(X, \tau, \mathsf{d}, \mathfrak{m})$ is an asymptotically Hilbertian extended metric-topological measure space with (X, d) complete, and if Ch denotes the associated Cheeger energy, defining $\mathcal{E} := \mathrm{Ch}$ one has that $(X, \mathscr{B}(\tau), \mathcal{E}, \mathfrak{m})$ is an energy measure space according to Definition 10.1. In addition:*

(a) $\Gamma(f) = |\mathrm{D}f|_w^2$ *for any $f \in \mathbb{V} \cap L^\infty(X, \mathfrak{m})$ and the formula*

$$\Gamma(f, g) := \lim_{\epsilon\downarrow 0} \frac{|\mathrm{D}(f + \epsilon g)|_w^2 - |\mathrm{D}f|_w^2}{2\epsilon} \qquad \text{in } L^1(X, \mathfrak{m})$$

extends the Γ operator from $\mathbb{V} \cap L^\infty(X, \mathfrak{m})$ to the whole of $\mathbb{V}$.

(b) *the extended distance*

$$\mathsf{d}_{\mathcal{E}}(x, y) := \sup\left\{ |f(x) - f(y)| : f \in \mathbb{V} \cap C_b(X), \ \Gamma(f) \leq 1 \right\} \tag{12.13}$$

satisfies $\mathsf{d}_{\mathcal{E}} \geq \mathsf{d}$ and, denoting by $\mathrm{Ch}_{\mathsf{d}_{\mathcal{E}}}$ the Cheeger energy associated to the new metric-topological structure $(X, \tau, \mathsf{d}_{\mathcal{E}}, \mathfrak{m})$, one has $\mathrm{Ch}_{\mathsf{d}_{\mathcal{E}}} = \mathcal{E}$.

(c) $\mathsf{d}_{\mathcal{E}} = \mathsf{d}$ *if and only if $f \in \mathbb{V} \cap C_b(X)$ and $|\mathrm{D}f|_w \leq 1$ $\mathfrak{m}$-a.e. in X implies $f \in \mathrm{Lip}_b(X, \tau, \mathsf{d})$ with $\mathrm{Lip}(f, \mathsf{d}) \leq 1$.*

Proof. By the locality and chain rule properties of $f \mapsto |\mathrm{D}f|_w$ stated in Proposition 6.3(c,e), the asymptotically Hilbertian assumption ensures that $\mathcal{E} := \mathrm{Ch}$ is a strongly local and Markovian Dirichlet form in $L^2(X, \mathfrak{m})$.

The proof of statement (a) can be obtained as in [4, Sec. 4], see in particular [4, Thm. 4.18]. For convenience we briefly sketch the proof. Let us set for $f, g \in \mathbb{V}$

$$G(f, g) := \lim_{\epsilon\downarrow 0} \frac{|\mathrm{D}(f + \epsilon g)|_w^2 - |\mathrm{D}f|_w^2}{2\epsilon} \qquad \text{in } L^1(X, \mathfrak{m}) .$$

Note that the limit above exists as a monotone limit by convexity of the minimal weak upper gradient, and that $\mathcal{E}(f,g) = \int G(f,g)\,\mathrm{d}\mathfrak{m}$. One first establishes the following chain rule for G: for all f, $g \in \mathbb{V}$ and $\phi : \mathbb{R} \to \mathbb{R}$ non-decreasing, Lipschitz and C^1, one has

$$G(f,\phi(g)) = \phi'(g)G(f,g)\,, \quad \int G(\phi(g),f)\,\mathrm{d}\mathfrak{m} = \int \phi'(g)G(g,f)\,\mathrm{d}\mathfrak{m}\,. \tag{12.14}$$

This is proven arguing as in [4, Lem. 4.7]. Moreover one follows [4, Prop. 4.17] using Lemma 12.7 to establish the Leibniz rule

$$\mathcal{E}(f,gh) = \int hG(f,g) + gG(f,h)\,\mathrm{d}\mathfrak{m} \tag{12.15}$$

for all f, g, $h \in \mathbb{V} \cap L^\infty(X,\mathfrak{m})$ with g, $h \geq 0$.

To prove the claim it is sufficient to show that G is bilinear and symmetric and therefore in turn it is sufficient to prove that $f \mapsto \int h|\mathrm{D}f|_w^2\,\mathrm{d}\mathfrak{m}$ is quadratic for every nonnegative bounded Borel h or, by approximation, $h \in \mathbb{V} \cap L^\infty(X,\mathfrak{m})$. Now use (12.14), (12.15) to write

$$\int h|\mathrm{D}f|_w^2\,\mathrm{d}\mathfrak{m} = \int hG(f,f)\,\mathrm{d}\mathfrak{m} = -\mathcal{E}(\tfrac{1}{2}f^2,h) + \mathcal{E}(f,fh)\,.$$

We conclude by quadraticity of the terms on the right hand side.

The proof of statement (b) is easy, since all functions $f \in \mathrm{Lip}_b(X,\tau,\mathsf{d})$ belong to $C_b(X)$ and $\sqrt{\Gamma(f)} = |\mathrm{D}f|_w \leq \mathrm{Lip}(f)$ $\mathfrak{m}$-a.e. in X. It follows that all the distances d_i approximating d from below are admissible in (12.13), so that $\mathsf{d}_\mathcal{E} \geq \mathsf{d}$. Since Theorem 9.2 ensures that $\mathcal{E} = \mathrm{Ch}$ is τ-upper regular, by Theorem 12.5 we obtain $\mathrm{Ch}_{\mathsf{d}_\mathcal{E}} = \mathcal{E}$.

In order to prove statement (c), notice that the inclusion $\{f \in C_b(X) \cap \mathbb{V} : |\mathrm{D}f|_w \leq 1\}$ in the class of 1-Lipschitz functions w.r.t. d implies, by the very definition of $\mathsf{d}_\mathcal{E}$, that $\mathsf{d}_\mathcal{E} \leq \mathsf{d}$. The converse is obvious, again by the definition of $\mathsf{d}_\mathcal{E}$. $\qquad\square$

13. EXAMPLES

Here we collect natural examples of energy measure spaces and extended metric measure spaces.

13.1. Degenerate Dirichlet forms.
Consider $X = \mathbb{R}^2$ equipped with the usual topology τ, the Borel σ-algebra $\mathcal{B}(\tau)$, and the standard Gaussian measure $\gamma(\mathrm{d}x) = (2\pi)^{-1}\mathrm{e}^{-|x|^2/2}\,\mathrm{d}x$. Consider further a Dirichlet form measuring energy only in the first coordinate, i.e.

$$\mathcal{E}(f) = \int |\partial_1 f(x_1,x_2)|^2\,\mathrm{d}\gamma(x_1,x_2)$$

for $f \in L^2(\mathbb{R}^2,\mathfrak{m})$ with $f(\cdot,x_2) \in W^{1,2}_{\mathrm{loc}}(\mathbb{R})$ for a.e. $x_2 \in \mathbb{R}$. Then $(\mathbb{R}^2,\mathcal{B}(\tau),\mathcal{E},\gamma)$ is an energy measure space according to Definition 10.1.

As the distance generated from the algebra $\mathcal{A}_\mathcal{E}^*$ according to (12.1d), (12.1e) one obtains

$$\mathsf{d}_\mathcal{E}\big((x_1,x_2),(y_1,y_2)\big) = \begin{cases} |x_1 - y_1| & \text{if } x_2 = y_2\,, \\ +\infty & \text{else} \end{cases}$$

and $(X,\tau,\mathsf{d}_\mathcal{E},\gamma)$ is an extended metric measure space according to Definition 4.7.

Note that the Bakry–Émery condition $\mathsf{BE}(1,\infty)$ is satisfied for this Dirichlet form, but $\mathcal{E}$ is not irreducible (see (10.17)); in fact, it is easy to construct measures with bounded densities but infinite $W_\mathcal{E}$ distance from $\mathfrak{m}$. Using a standard approximation by restriction, one can check that $\mathcal{E}$ is τ-upper regular. It follows that $\mathcal{E}$ coincides with the Cheeger energy induced by $\mathsf{d}_\mathcal{E}$, by Theorem 12.5.

13.2. **Abstract Wiener spaces.** Consider a separable Banach space X (or, more generally, a locally convex topological space) and a centered, non-degenerate Gaussian measure γ in X. The Cameron-Martin space $\mathscr{H} \subset X$ is the image under the mapping

$$Rf := \int f(x)x \, \mathrm{d}\gamma(x) \qquad f \in L^2(X, \gamma)$$

of the so-called reproducing kernel of γ, namely the closure H in $L^2(X, \gamma)$ of $\{\langle x', x\rangle\}_{x' \in X'}$ (here X' is the topological dual of X and $\langle \cdot, \cdot \rangle$ is the canonical pairing). It is easily seen that $R : H \to \mathscr{H}$ is injective, hence $\mathscr{H}$ inherits from $L^2(X, \gamma)$ a Hilbert structure. The structure $(X, \gamma, \mathscr{H})$ is called abstract Wiener space.

Denote by $\mathscr{H}_b^\infty$ the set of smooth, bounded cylinder functions, i.e. the functions f of the form $f(x) = \phi(\langle x_1', x\rangle, \ldots, \langle x_n', x\rangle)$, for $x_1', \ldots, x_n' \in X'$ and $\phi : \mathbb{R}^n \to \mathbb{R}$ smooth and bounded. For such a cylinder function we define its gradient via

$$\nabla_{\mathscr{H}} f(x) = \sum_{i=1}^n \frac{\partial \phi}{\partial z_i}(\langle x_1', x\rangle, \ldots, \langle x_n', x\rangle) R(\langle x_i', \cdot\rangle) \, .$$

It is well-known that the quadratic form

$$\mathcal{E}(f) = \int |\nabla_{\mathscr{H}} f|_{\mathscr{H}}^2 \, \mathrm{d}\gamma \qquad f \in \mathscr{H}_b^\infty$$

is closable, that $\mathscr{H}_b^\infty$ is dense in $L^2(X, \gamma)$ and that $\mathcal{E}$ admits a carré du champ operator, see [15]. Thus $(X, \mathcal{B}(\tau), \mathcal{E}, \gamma)$, with τ being the weak or strong topology of the Banach space X is an energy measure space according to Definition 10.1.

The associated semigroup P is given by Mehler's formula

$$\mathsf{P}_t f(x) = \int f(\mathrm{e}^{-t}x + \sqrt{1 - \mathrm{e}^{-2t}}y) \, \mathrm{d}\gamma(y) \, .$$

From this one can check that the Bakry–Émery condition $\mathsf{BE}(1, \infty)$ holds. Using a Rademacher type theorem for abstract Wiener spaces, see [20], one can check that the induced distance according to (12.1d), (12.1e) is the Cameron–Martin distance

$$\mathsf{d}_{\mathcal{E}}(x, y) = \begin{cases} |x - y|_{\mathscr{H}} & \text{if } x - y \in \mathscr{H} \, , \\ +\infty & \text{else} \, . \end{cases}$$

Thus $(X, \tau, \mathsf{d}_{\mathcal{E}}, \gamma)$ is an extended metric measure space according to Definition 4.7.

13.3. **Configuration spaces.** Configuration spaces appear naturally as the state space for systems of infinitely many indistinguishable diffusing particles. Let M be a Riemannian manifold with metric tensor $\langle \cdot, \cdot \rangle$. The configuration space Υ over M is the set of all locally finite counting measures, i.e.

$$\Upsilon = \{\gamma \in \mathcal{M}(M) : \gamma(K) \in \mathbb{N}_0 \ \forall K \subset M \text{ compact}\} \, .$$

The space Υ is equipped with the vague topology, denoted by τ, in duality with continuous and compactly supported functions.

A natural differentiable and energy structure on the configuration space has been introduced in [1], by lifting the geometry on M, as we shall briefly describe. The tangent space $T_\gamma \Upsilon$ consists of all γ-square integrable vector fields on M and is equipped with the inner product

$$\langle V_1, V_2 \rangle_\gamma = \int \langle V_1(x), V_2(x) \rangle_x \, \mathrm{d}\gamma(x) \, .$$

Let Cyl_b^∞ be the set of smooth and bounded cylinder functions, i.e. functions $F : \Upsilon \to \mathbb{R}$ of the form $F(\gamma) = g(\gamma(\phi_1), \ldots, \gamma(\phi_n))$ where $g \in C_b^\infty(\mathbb{R}^n)$ and $\phi_1, \ldots, \phi_n \in C_c^\infty(M)$ and we write $\gamma(\phi) = \int \phi \, \mathrm{d}\gamma$. Given $F \in Cyl_b^\infty$ we define its gradient at γ as the vector field on M given by

$$T_\gamma \Upsilon \ni \nabla^\Upsilon F(\gamma) = \sum_{i=1}^n \frac{\partial g}{\partial z_i}(\gamma(\phi_1), \ldots, \gamma(\phi_n)) \nabla \phi_i \, .$$

Similarly, for a cylindrical "vector field" on Υ of the form $W = \sum_i F_i V_i$ with $F_i \in Cyl_b^\infty$ and V_i smooth, compactly supported vector fields on M, we define its divergence as

$$\mathrm{div}^\Upsilon(W)(\gamma) = \sum_i \langle \nabla^\Upsilon F_i, V_i \rangle_\gamma + F_i(\gamma)\gamma(\mathrm{div} V_i) \ .$$

The natural reference measure on Υ is the Poisson measure π, that can be characterized by its Laplace transform, i.e. for all $f \in C_b(M)$:

$$\int_\Upsilon \exp\big(\gamma(f)\big) \, \mathrm{d}\pi(\gamma) = \exp\left(\int_M \exp\big(f(x)\big) - 1 \, \mathrm{dvol}(x)\right) \ .$$

This is (up to the intensity) the unique probability measure such that the gradient and the divergence are adjoint in $L^2(\Upsilon, \pi)$, see [1, Thm. 3.2]. The quadratic form

$$\mathcal{E}(F,F) = \int \langle \nabla^\Upsilon F, \nabla^\Upsilon \rangle \, \mathrm{d}\pi$$

defined on Cyl_b^∞ is closable to a Dirichlet form admitting a carré du champ operator, see [1, Cor. 1.4], [38, Prop. 1.4], so that $(\Upsilon, \mathcal{B}(\tau), \mathcal{E}, \pi)$ is a Energy measure space according to Definition 10.1. The associated semigroup is given by the evolution of infinitely many independent Brownian particles on M.

The induced distance according to (12.1d), (12.1e) is known to be the L^2-transport distance between (non-normalized) configurations (see [38, Thm. 1.5]), i.e.

$$\mathsf{d}_\mathcal{E}(\gamma, \eta) = \inf_q \sqrt{\int \mathsf{d}^2(x,y) \, \mathrm{d}q(x,y)} \ ,$$

where d is the Riemannian distance and the infimum is over all couplings q of γ and η. $(X, \tau, \mathsf{d}_\varepsilon, \pi)$ is now an extended metric measure space according to Definition 4.7.

It is shown in [21, Prop. 2.3] that $\mathcal{E}$ coincides with the Cheeger energy constructed from $d_\mathcal{E}$ (as defined in [3] based on Lipschitz constants, but similar arguments apply to the construction used here based on asymptotic Lipschitz constants).

A detailed study of curvature bounds for configuration spaces can be found in [21]. It has been shown that various notions of curvature bounds lift from the base space M to the configuration space. In particular, if the Ricci curvature of M is bounded below by K, an Evolution Variational Inequality and the Bakry–Émery gradient estimate with constant K for the semigroup hold on the configuration space.

References

[1] S. Albeverio, Y. Kondratiev, and M. Röckner. *Analysis and geometry on configuration spaces*. J. Funct. Anal., **154** (1998), 444–500. 1998.

[2] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2008.

[3] L. Ambrosio, N. Gigli, and G. Savaré. *Calculus and heat flow in metric measure spaces and applications to spaces with Ricci bounds from below*. Invent. Math., **195** (2014), 289–391.

[4] L. Ambrosio, N. Gigli, and G. Savaré. *Metric measure spaces with Riemannian Ricci curvature bounded from below*. Duke Math. J., **163** (2014), 1405–1490.

[5] L. Ambrosio, N. Gigli, and G. Savaré. *Bakry-Émery curvature-dimension condition and Riemannian Ricci curvature bounds*. Annals of Probability, **43** (2015), 339–404.

[6] L. Ambrosio, M. Colombo, and S. DiMarino. *Sobolev spaces in metric measure spaces: reflexivity and lower semicontinuity of slope*. Advanced Studies in Pure Mathematics, **67** (2015), 1–58.

[7] L. Ambrosio, A. Mondino, G. Savaré. *On the Bakry-Émery condition, the gradient estimates and the Local-to-Global property of RCD*(K,N) metric measure spaces*. The Journal of Geometric Analysis, to appear. ArXiv e-prints 1309.4664 (2013) 1–29.

[8] L. Ambrosio, M. Erbar, and G. Savaré. In preparation.

[9] L. Ambrosio and D. Trevisan. *Well posedness of Lagrangian flows and continuity equations in metric measure spaces*. Analysis & PDE, **7** (2014), 1179–1234.

[10] D. Bakry. *Functional inequalities for Markov semigroups*, in Probability measures on groups: recent directions and trends, Tata Inst. Fund. Res., Mumbai, 2006, pp. 91–147.

[11] D. Bakry and M. Émery. *Diffusions hypercontractives,* in Séminaire de probabilités, XIX, 1983/84, vol. 1123, Springer, Berlin, 1985, pp. 177–206.

[12] D. Bakry, I. Gentil, and M. Ledoux. *Analysis and Geometry of Markov Diffusion Operators.* Vol. 348 of Grundlehren der mathematischen Wissenschaften, Springer, 2014.

[13] M. Biroli and U. Mosco. *A Saint-Venant type principle for Dirichlet forms on discontinuous media.* Ann. Mat. Pura Appl. (4), **169** (1995), 125–181.

[14] V.I. Bogachev. *Measure Theory I, II.* Springer, Berlin, 2007.

[15] N. Bouleau and F. Hirsch. *Dirichlet forms and analysis on Wiener spaces.* Vol. 14 of De Gruyter studies in Mathematics, De Gruyter, 1991.

[16] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert.* North-Holland Mathematics Studies, No. 5. Notas de Matemática (50), North-Holland Publishing Co., 1973.

[17] J. Cheeger. *Differentiability of Lipschitz functions on metric measure spaces.* Geom. Funct. Anal., **9** (1999), 428–517.

[18] S. Daneri and G. Savaré. *Lecture notes on gradient flows and optimal transport,* in Optimal Transportation. Theory and Applications, Y. Ollivier, H. Pajot, C. Villani eds. London Mathematical Society Lecture Note Series **413** (2014), 100–144.

[19] S. Daneri and G. Savaré. *Eulerian calculus for the displacement convexity in the Wasserstein distance,* SIAM J. Math. Anal., **40** (2008), 1104–1122.

[20] O. Enchev and D. Stroock. *Rademacher's theorem for Wiener functionals,* Ann. Probab., **21** (1993), 25–33.

[21] M. Erbar, and M. Huesmann. *Curvature bounds for configuration spaces.* Calc. Var. Partial Differential Equations, DOI:10.1007/s00526-014-0790-1, (2014).

[22] M. Erbar, K. Kuwada and K.T. Sturm. *On the equivalence of the entropic curvature-dimension condition and Bochner's inequality on metric measure spaces.* Invent. Math., DOI: 10.1007/s00222-014-0563-7, (2014).

[23] M. Fukushima, Y. Oshima, and M. Takeda. *Dirichlet forms and symmetric Markov processes.* Vol. 19 of de Gruyter Studies in Mathematics, Walter de Gruyter & Co., Berlin, extended ed., 2011.

[24] N. Gigli. *On the heat flow on metric measure spaces: existence, uniqueness and stability.* Calc. Var. Partial Differential Equations, **39** (2010), 101–120.

[25] N. Gigli, A. Mondino, and G. Savaré. *Convergence of pointed non-compact metric measure spaces and stability of Ricci curvature bounds and heat flows.* arXiv:1311.4907 (2013).

[26] N. Gigli. *On the differential structure of metric measure spaces and applications.* Memoirs of the AMS, **236**, n. 1113 (2014).

[27] N. Gigli and H.Bangxian. *The continuity equation on metric measure spaces.* Calc. Var. Partial Differential Equations, **53** (2015), 149–177.

[28] M. Hino, J.A. Ramirez. *Small-time Gaussian behavior of symmetric diffusion semigroups.* The Annals of Probability, **31** (2003), 1254–1295.

[29] H.G. Kellerer. *Duality theorems for marginal problems.* Z. Wahrsch. Verw. Gebiete, **67** (1984), 399–432.

[30] P. Koskela and Y.Zhou. *Geometry and Analysis of Dirichlet forms.* Adv. Math., **231** (2012), 2755–2801.

[31] P. Koskela, Y. Zhou, and N.Shanmugalingam. *Geometry and Analysis of Dirichlet forms II.* J. Funct. Anal., **267** (2014), 2437–2477.

[32] K. Kuwada. *Duality on gradient estimates and Wasserstein controls.* Journal of Functional Analysis, **258** (2010), 3758–3774.

[33] J.L. Lions and E. Magenes. *Non Homogeneous Boundary Value Problems and Applications,* Vol. I. Springer, New-York, 1972.

[34] S. Lisini. *Absolutely continuous curves in extended Wasserstein–Orlicz spaces.* ArXiv e-prints 1402.7328 (2014), 1–22.

[35] J.R. Munkres *Topology: a first course.* Prentice-Hall, Inc., Englewood Cliffs, N.J., 2000.

[36] S.I. Ohta and K.T. Sturm. *Non-contraction of heat flow on Minkowski spaces.* Arch. Ration. Mech. Anal., **204** (2012), 917–944.

[37] F. Otto and C. Villani. *Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality.* J. Funct. Anal., **173** (2000), pp. 361–400.

[38] M. Röckner and A. Schied. *A general duality theorem for marginal problems.* J. Funct. Anal., **169** (1999), 325–356.

[39] L. Schwartz. *Radon measures on arbitrary topological spaces and cylindrical measures,* Tata Institute of Fundamental Research Studies in Mathematics, No. 6, 1973.

[40] S. Serfaty. *Γ-convergence of gradient flows and applications to Ginzburg-Landau vortex dynamics.* Topics on concentration phenomena and problems with multiple scales, 267–292, Lect. Notes Unione Mat. Ital., 2, Springer, Berlin, 2006.

[41] T. Shioya. *Metric measure geometry – Gromov's theory of convergence and concentration of metrics and measures.* IRMA series of the European Mathematical Society, *to appear.* ArXiv e-prints 1410.0428 (2014), 1–172.

[42] P. Stollmann. *A dual characterization of length spaces with application to Dirichlet metric spaces.* Polska Akademia Nauk. Instytut Matematyczny. Studia Mathematica, **198** (2010), 221–233.

[43] K.T. Sturm. *Is a diffusion process determined by its intrinsic metric?* Chaos Solitons Fractals, **8** (1997), 1855–1860.

[44] K.T. Sturm. *The space of spaces: curvature bounds and gradient flows on the space of metric measure spaces.* ArXiv e-prints 1208.0434 (2012), 1–75.

[45] C. Villani. *Topics in optimal transportation.* Graduate Studies in Mathematics, **58**, American Mathematical Society, 2003.

[46] C. Villani. *Optimal transport. Old and new.* Grundlehren der Mathematischen Wissenschaften **338**, Springer-Verlag, Berlin, 2009.

[47] F.Y. Wang. *Equivalent semigroup properties for the curvature-dimension condition.* Bull. Sci. Math., **135** (2011), 803–815.

[48] N. Weaver. *Lipschitz Algebras.* World Scientific, 1999.

[49] N. Weaver. *Quantum relations.* Mem. Amer. Math. Soc., **215** (2012), 81–140.

[50] D. Zaev. L^p *Wasserstein distances on state and quasi-state spaces of C^* algebras.* ArXiv 1505.06061.

SCUOLA NORMALE SUPERIORE, PISA
E-mail address: `luigi.ambrosio@sns.it`

UNIVERSITY OF BONN
E-mail address: `erbar@iam.uni-bonn.de`

PAVIA UNIVERSITY
E-mail address: `giuseppe.savare@unipv.it`